APPENDIX A

Top Level System Requirements Document

(Last Edited - March 23, 2010) Lorenzo Allori and Marcello Botrugno

1. I	INTRODUCTION	4
1.1	Purpose	4
1.2	Intended Audience	4
1.3	Project Scope	5
1.4	References	5
2. (OVERALL DESCRIPTION	6
2.1	Product Perspective	6
2.2	Modular Design	6
2.3	Product Features	7
2.4 2.4.1 2.4.2 2.4.3 2.4.4 2.4.5 2.4.6 2.4.7 2.4.8 2.4.9	On-Site Fellows	
2.52.6	Design and Implementation Constraints User and Developer Documentation	
	SYSTEM FEATURES	
3.1.1 3.1.2	r	13
3.2.1 3.2.2	r · · ·	13
3.3 3.3.1 3.3.2	1	15

3.4	Logging user actions and data revisions	
3.4.1	Description	
3.4.2	Functional Requirements	16
3.5	Document Digitization	17
3.5.1	Description	
3.5.2	Functional Requirements	17
3.6	System Administration	18
3.6.1	Description	
3.6.2	Functional Requirements	18
4. E	XTERNAL INTERFACE REQUIREMENTS	19
4.1	Hardware Interfaces	10
4.1.1	Digitization hardware	
4.1.2	O .	
5. O	THER NONFUNCTIONAL REQUIREMENTS	20
5.1	Performance Requirements	20
5.2	Security Requirements	20
5.3	Software Quality Attributes	
5.3.1	Usability	
5.3.2	Cross-platform compatibility	
5.3.3	Reliability	21
6. O	THER REQUIREMENTS	22
6.1	Data migration	22
U.I	~ www 11115 1 with 11 11 11 11 11 11 11 11 11 11 11 11 11	

1. Introduction

1.1 Purpose

The Medici Archive Project's mission is to bring the unparalleled resources of the Medici Granducal Archive for the Arts and Humanities to a broad international public for the first time, by way of the latest information technology.

Therefore, after eleven years, the Medici Archive Project (hereafter referred to as MAP) needs to develop more efficient and productive software system using up-to-date technologies, giving the research Fellows the possibility to perform many new tasks.

This document contains the outline of the first phase of the software planning.

1.2 Intended Audience

This document contains the software's functional and non-functional requirements. It is intended to be read by the MAP Staff, Project Managers, and Software Developers.

Chapter 2 gives a general description of the system. It is intended to be read by the MAP staff.

Chapter 3 contains analysis of both functional and non-functional requirements in full detail. This chapter will be evaluated by the MAP staff, and by the Software Architects. It will be used in order to create the Top Level System Design Document in the second phase of the software planning.

1.3 Project Scope

The new system will offer more functions than the current one, making the work of the Fellows increasingly efficient and productive. By providing access to digitized primary sources, MAP will be able to hire Distance Fellows who will work on the software from remote locations. This way MAP will be able to reduce costs while increasing productivity. Additionally, the system will create a link between the MAP and the outside world, expanding a community of scholars of early modern history. These scholars, both professional academics and graduate students, will be able to comment and discuss the data interactively while consulting the database. Therefore, the system must be web-based in order to be accessible from the Internet.

1.4 References

```
Minutes of meeting:
```

Minutes of Meeting 1 – Functional requirements

Minutes of Meeting 2 – Digitization process requirements

Minutes of Meeting 3 – State Archive in Florence requirements

Minutes of Meeting 3 – Catalog Description

Minutes of Meeting 4 – Community requirements

Minutes of Meeting 5 – History log module requirements

Current websites:

http://documents.medici.org (MAP online read-only database) http://courses.medici.org (MAP online teaching software)

2. Overall Description

2.1 Product Perspective

This system will be an enhanced version of the Microsoft Access-based, which MAP Fellows currently use as their data-entry software.

Unlike the current software, the new system will be easily accessible off-site, and therefore, must be fully accessible from the Internet. To achieve this, the functionalities of the current software must be translated and adapted in order to comply with those of a web-based system.

2.2 Modular Design

The structure of the new system will be modular. This means that the software is divided into several modules, which fulfill different functional tasks.

The access to the new system will be carried out through an authentication phase; a user can access certain parts of the application based on his group credentials. An *Authentication module* regulates this process.

A *History Log module*, which tracks the users' actions in the system, will be connected to the *Authentication module*.

The *Data Entry module* is defined by the part of the system where the documents (accompanied by digitized images), historic people, and historic places are stored. The Fellows will be able to read, create, and modify data; the scholarly community will only be able to read the documents, and not to modify or create them.

The documents module is divided into four sections:

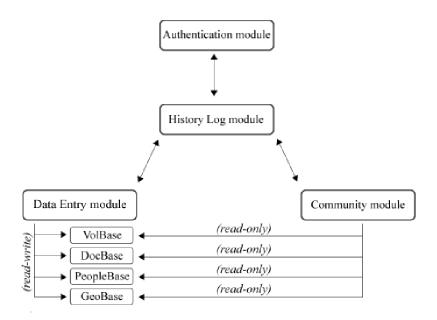
VolBase (contains the digitized volumes of the Medici Granducal Archival Collection)

DocBase (stores identifying and indexing information regarding the documents)

PeopleBase (stores biographical information regarding the individuals and corporate bodies whose names appear in the DocBase)

GeoBase (contains the hierarchy of places used in the DocBase from continents and nations down to individual buildings, institutions and sites)

The *Community module* will manage all of the online scholarly community's comments, questions, and answers regarding each item present in the database (volumes, documents, historic people, historic places).



2.3 Product Features

System features requirements list:

- 1. Login to the system with username and password
- 2. Creation and/or update and/or deletion of database records¹:
 - Flip-through digitized documents in a volume
 - Create/update/delete volumes records
 - Create/update/delete documents records
 - View digitized documents
 - Create/update/delete bio pages records
 - Create/ update/delete places records
 - Send personal messages to other users of the system
- 3. Search records:
 - Search through the volumes, document transcriptions, digitized documents, bios and places records
- 4. Community actions:

¹ A database record is one entry in the database

- Post comments/questions/answers on volumes, document transcriptions, bios and places records
- Send personal messages to the other users of the system
- Edit/view personal profiles
- 5. Moderate community comments/questions/answers:
 - View/evaluate/delete community comments/questions/answers
- 6. Log tracking of users' actions in the system:
 - System access
 - New record creation
 - Record deletion
 - Record modification
- 7. Document digitization:
 - Upload digitized images to the database
 - Verify digitized documents
 - Write catalog descriptions
- 8. System Administration:
 - Manage users
 - Read system and user actions logs
 - Read database item review history
 - Backup and restore system data and users profiles

2.4 User Groups

These are the defined user groups:

- Administrators
- On-Site Fellows
- Senior Distance Fellows
- Distance Fellows
- Digitization Technicians
- Community Users
- Guests

These are "special roles" which can be assigned to one or more users:

- Distant Fellows Coordinator
- Community Coordinator
- Digitization Coordinator

2.4.1 Administrators

The Administrators perform all the administration tasks:

- Create/delete/modify users and user-profiles
- Read system and user logs
- Read database items review history
- Backup and/or restore data

2.4.2 On-Site Fellows

The On-Site Fellows work from September to July, four days a week, six hours a day. They begin by selecting one of the 6,429 volumes relevant to a particular topic or time period of special interest. They contextualize the history and physical condition of the volume; describe and appraise its codicological properties; read and evaluate all the letters and other documents in the volume; select those that appear most pertinent to the historical database; transcribe the original text (mostly in Italian, but also in Latin, Spanish, French, German, and Hebrew); decipher abbreviations and make sense of linguistic impasses such as lacunae, misspellings, deteriorated text, ciphered words and coded expressions. Finally, they write a succinct yet exhaustive synopsis of the specific documentary extract, synthesizing factual information, evaluating the plausibility of statements made in the document, and proposing solutions to the meaning of the document and its relation to wider historical contexts. Their ability to synthesize historical information as well as their uniquely privileged contact with the actual documents authorizes them to check and review (or "vet") the data entered by other users.

For performing special actions, the On-Site Fellows will be assigned (or "flagged") as Distance Fellows Coordinators, Community Coordinators and/or Digitization Coordinators [see 2.4.8].

2.4.3 Senior Distance Fellows

The Senior Distance Fellows will work off-site on previously digitized and uploaded documents. The users that belong to this category are allowed to create and modify any of their own entries on documents; additionally they will be able to vet the Distance Fellows' work.

2.4.4 Distance Fellows

The Distance Fellows will work off-site on previously digitized and uploaded documents. The users that belong to this category are allowed to create and modify and delete any of their own entries on documents, volumes, bios or places, however a Distance Fellow cannot

delete or vet any other type of content. The Research Director² assigns the vetting of the documents entered by the Distance Fellows to one or more On-Site Fellows or Senior Distance Fellows.

2.4.5 Community Users

Community users can perform any kind of search and they can post comments/questions/answers on each item that matches their search criteria. They will be able to see the digitized documents by flipping through the volumes.

A link to our web-based online course system (http://courses.medici.org] will be present to educate the community users on our current curriculum offerings; we expect to recruit many interested and qualified future online scholars through this channel.

2.4.6 Digitization Technicians

These users are in charge of the document digitization. Digitization Technicians (DTs] will log in to the system, upload the digitized documents and edit the catalog descriptions [see *Minutes of Meeting 3 – State Archive in Florence requirements*]. DT's do not have any other privileges.

2.4.7 Guests

Users that browse our system without registering a username and password belong to this group. They will have access to the database to perform any kind of search but access to digitized documents will be forbidden. They will be able to apply for a username and password in order to become regular Community Users. In order to join the MAP Community they will need to fill out a registration form and accept a community rules disclaimer.

2.4.8 Special Roles

• *Distance Fellows Coordinator (DFC)*: this is a "flag" that an On-Site Fellow or a Senior Distance Fellow can obtain. The DFC is responsible for monitoring the Distance Fellows' work. The DFC is notified of any new records that Distance Fellows create, but is not responsible for vetting them.

• Community Coordinator (CC): this is a "flag" that an On-Site Fellow or a Community User can obtain, with the responsibility of evaluating the comments posted by the Community Users. A CC will be notified by the system when a comment is posted, after which she will review it and decide if it matches the community disclaimer rules. If not, the comments will not be approved and will be deleted. A CC will also be able to close a conversation thread.

10

² The Research Director is an On-Site Fellow, who is responsible for the whole Medici Archive Project's research on the Medici Granducal Archive Collection.

• *Digitization Coordinator (DC)*: this is a "flag" that an On-Site Fellow can obtain, with the responsibility of monitoring and quality-assurance of document digitization. The system will notify the DCs about new documents that DTs upload to the system.

2.4.9 Workflow

The workflow is divided into two steps; a defined class of users performs the following actions:

1st step – Preliminary reviewing

Digitize documents: Digitization Technician, Digitization Coordinator

1st reading (data entry): On-Site Fellows, Distance Fellows

2nd reading (pre-vetting comments): Community Users

3rd reading (preliminary vetting): On-Site Fellows, Senior Distance Fellows

2nd step - Continuous reviewing

1st review (community comments): Community Users

2nd review (answer comment): On-Site Fellows, Senior Distance Fellows and Distance Fellows

(This process continues until a Community Coordinator decides to close the thread)

2.5 Design and Implementation Constraints

The graphic user interface has to be web-based.

Lightweight production of web pages is desirable to guarantee fast, reliable access to the system.

It is acceptable to use an opensource RDBMS.

2.6 User and Developer Documentation

A *User Manual* will be produced.

Video lessons will be created to get new users acquainted with the system.

In addition, a *Developer Manual* will be produced.

3. System Features

3.1 Access to the System – Authentication Module

3.1.1 Description

The login to the system is performed through username and password. All the users must be registered except the guest users [see 2.4.7].

After the login phase, every user has access to some specific part of the systems based on its group credentials [see 2.4].

3.1.2 Functional Requirements

3.1.2.1 **REQ-01 – Authentication methods**

Authentication methods must be implemented in accordance with the more recent standards for web authentication.

3.2 Data Entry – Documents Module

3.2.1 Description

Data entry is one of the most important tasks. The Documents Module is the "workbench" used by Fellows to index documents.

To review all the features needed, see the Appendix D - MAPDocSources V4D manual.

To review what has been discussed during the meeting with the Fellows, check the document *Minutes of Meeting 1 – Functional requirements*.

These features will be discussed and analyzed in depth in Phase 2, the System Design Phase [see *Appendix B, Top Level System Design Document*].

3.2.2 Functional Requirements

3.2.2.1 REQ-02 – View digitized documents

Scholars will be able to see and transcribe digitized manuscripts; therefore a web-based tool (a manuscript viewer) must be created and integrated. During Phase 2, the Design Phase, MAP will investigate how to accomplish this requirement.

3.2.2.2 REQ-03 – Browsing through digitized documents in a volume

The goal is to achieve a realistic "flip-through" effect to be utilized on the digitized documents of a virtual volume. This "flipping" will have to be as similar to flipping through a real volume as possible. If the volume contains inserts it will be possible to see those inserts and flip through them.

In the virtual volume, documents must have a resolution high enough to be easily read by the users.

3.2.2.3 REQ-04 – Creating/modifying a volume in the VolBase

Creating or modifying a volume requires the Fellow to filling in of a large amount of information (volume description, its correspondents, and so on).

These tasks have to be done without the web page being refreshed. During Phase 2, a web technology will be chosen to accomplish this requirement.

3.2.2.4 REQ-05 – Creating/modifying a document in the DocBase

Creating or modifying a document requires the following steps:

- Identify a document relevant to the MAP topics list
- Identify the sender and recipient that, if not present in the database, must be inserted through the bio data entry section [see 3.2.2.5]. It is required to see if that person already exists in the database or not. For this reason, an "autocomplete" dropdown text field is a must
- Identify the sender and recipient locations that, if not present in the database, will be added in the places data entry section [see 3.2.2.6]. It is required to see if that person already exists in the database or not. For this reason, an "autocomplete" dropdown text field is a must
- Add the document transcription in the original language
- Add the transcription synopsis in English
- Link the document to the MAP topics it refers to

These tasks must be done without the web page being refreshed. During Phase 2, a web technology will be chosen to accomplish this requirement.

3.2.2.5 REQ-06 – Creating/modifying a bio in the PeopleBase

During the Fellows' work on the documents it is crucially important to add the sender and the recipient name. Those historic figures must be present in the PeopleBase, if not it is necessary to add them.

When adding a new bio, the following fields will be filled in: vital statistics, name variants, parents, children, spouses, roles and occupations. In case of children, parents and spouses, the person must be linked to the pertinent other bios.

3.2.2.6 REQ-07 – Creating/modifying a place in the GeoBase

The database contains the records of all the places (if recognizable) contained in the documents. These places can refer to modern nations, cities that exist now or that do not exist anymore, as well as churches, villas or castles.

Each Place is indexed in the GeoBase according to its modern name, in the principal language spoken in that country. For example, the Belgian city of Antwerp is searchable by entering its Flemish name, Antwerpen.

Google maps must be integrated in the system to show clearly where the location is.

3.3 Community Module

3.3.1 Description

Though read-only browsing in our database is free and is a chief prerogative of the whole system, users belonging to the Guest group [see 2.4.7] are forbidden from browsing digitized documents, and from posting comments.

Registered Community users are authorized to see digitized documents and to post comments, questions and answers [see 2.4.5].

The comments, questions and answers entered by the community have their own life cycle. If a user posts a comment or a question, another user can answer it. The discussion will go on until the Community Coordinator [see 2.4.8] decides to close the discussion thread.

The Community users will be able to send personal messages to each other. The ability to communicate this way will be very important in establishing relationships between scholars who have interests in common.

3.3.2 Functional Requirements

3.3.2.1 REQ-08 – New user registration form

New users will be able to ask for a username and password through a registration form. A community rules disclaimer will be created, one that new users must agree to, in order to successfully register.

3.3.2.2 REQ-09 – Search the database

When users log in, they will be able to browse the data using a search mask similar to the one we have at our current on-line database at http://documents.medici.org. Search results will include digitized documents as well.

3.3.2.3 REQ-10 – Posting comments, questions and answers

Community users will be able to post comments, questions and answers, according to the community rules disclaimer. If the comments posted do not comply with the community rules, the Community Coordinators [see 2.4.8] will delete them. The Community Coordinators will be able to close a discussion thread blocking any further comments.

3.4 Logging user actions and data revisions

3.4.1 Description

Every modification to the data in the system must be logged and saved. This is needed for monitoring purposes and to track user-actions to see if the users are correctly operating the interface.

When a system error occurs these logs will be reviewed in order to discover the reason for that failure.

To restore any type of data loss, and as an enhancement to regularly scheduled backup procedures, a history of revisions of every single item record (volume, document, bio, place) present in the database must be available to the Administrators user group [see 2.4.1].

3.4.2 Functional Requirements

3.4.2.1 REQ-11 – Access to the system

Registered users access the system by entering a username and a password. Every login will be recorded, including the following fields: username, IP address, date and time.

3.4.2.2 REQ-12 – Adding new data

The creation of new records will be logged, tracing username, IP address, date and time, as well as a description of the user action performed (for example "added sender into doc id 2305" or "added role and occupation to person id 34283"). A copy of that item with date and time will be stored in the database for the history of revisions.

3.4.2.3 REQ-13 – Deleting data

The deletion of new records will be logged tracing username, IP address, date and time, as well as a description of the user action performed (for example "deleted place record id 4442" or "deleted doc id 2308"). At the time of the deletion, a copy of the deleted item with date and time will be stored in the database for the history of revisions.

3.4.2.4 REQ-14 – Modifying data

Every modified item is logged tracing username, IP address, date and time, as well as a description of the user action performed (for example "modified field recipient in doc id 3256" or "modified topic in doc id 4492"). A copy of the modified item with date and time is stored for the history of revisions.

3.5 Document Digitization

3.5.1 Description

Document digitization is a high priority requirement in order to work off-site; it also helps preserve the precious original documents from wear and tear.

The first twenty volumes of the Medici Granducal Archival Collection have already been digitized and will be uploaded to the system.

All subsequent volumes will be digitized in their entirety and will be uploaded one at a time to the new system.

3.5.2 Functional Requirements

3.5.2.1 REQ-15 – Upload digitized documents to the database

A standalone program or a web interface will be given to the Digitization Technician in order to upload the digitized images to the system in the correct format.

Digitized documents can be stored in the file system or in the database. This decision will be made in the System Design Phase (Phase 2).

3.5.2.2 REQ-16 – Digitization process

The Digitization Technician will follow the State Archive's cataloguing conventions [see *Minutes of Meeting 3 – State Archive in Florence requirements*]. The reproduction will be made using digital camera technology [see 4.1.1.]

Once the digitization of a volume is completed, the Digitization Technician will upload the images to the system.

The Digitization Coordinator [see 2.4.8] will check the quality of the images produced, in order to guarantee an easy reading for the Fellows.

3.6 System Administration

3.6.1 Description

Administration tasks are performed through a particular part of the web interface that is accessible only to the Administrators user group [see 2.4.1].

3.6.2 Functional Requirements

3.6.2.1 REQ-17 – User accounts management

It will be possible to perform these tasks:

- Show the user list
- Create new users
- Modify user accounts
- Delete users

3.6.2.2 REQ-19 – Backup and restore

A backup and restore option will be available.

3.6.2.3 **REQ-20 – Logs review**

A mask for searching through logs will be available.

Logs will be searched by each relevant field. Furthermore, it would be possible to cross-search these items: user, IP address, action, information, date and time.

3.6.2.4 REQ-21 – Revisions review

A history review option will be available to Administrators by clicking a "review history list" link, displayed on each volume, document, bio or place page.

4. External Interface Requirements

4.1 Hardware Interfaces

4.1.1 Digitization hardware

The digitization will be performed using a digital camera taking pictures at minimum 300 DPI.

4.1.2 Servers' Hardware

The servers' hardware will guarantee enough storage space for the digitized document images, and enough speed for the performance required. Technical specifications will be investigated in the System Design Phase (Phase 2).

5. Other Nonfunctional Requirements

5.1 Performance Requirements

The *response time---* i.e. the time a system or functional unit takes to react to a given input---of the new system will be faster than current software. The *response time* of the web application used to view the digitized documents must be reasonable in order to provide fast and reliable access to the digitized manuscripts.

5.2 Security Requirements

The data will be protected from web attacks.

Read-only access to the system is open and free, meaning that everyone can access it from the Internet. To see digitized documents or to participate in the community discussions, a user must register to get a username and password.

The write-access to the database is permitted only to these user groups: Administrators, On-Site Fellows, Senior Distant Fellows, and Distant Fellows.

5.3 Software Quality Attributes

5.3.1 Usability

The entry data interface will be usable and accessible to Fellows. This means that all the tasks the Fellows perform will be executed as fast as possible, using the fewest possible mouse clicks.

Response time of the search tasks performed by Fellows and Community users will be reduced to a minimum.

The possibility of refining search results is necessary.

5.3.2 Cross-platform compatibility

The system must be compatible with Apple, Microsoft, and Linux operating systems.

5.3.3 Reliability

The software reliability is an important requirement. Any data loss will be avoided.

6. Other Requirements

6.1 Data migration

Volumes, documents, bios and places records reside now in the current software. These data will be transferred into the new system, which will be using a RDBMS that is compatible with a web-based system.

A web-based read-only system (http://documents.medici.org) has already been developed and it is based on an opensource database, MySQL. We do have tools to migrate the vetted data from the current software (Microsoft Access) to MySQL.

The migration task has to be planned and verified to avoid data loss, and will be investigated in the System Design Phase (Phase 2).

Minutes of Meeting

Minutes of Meeting 1 MAP Software Redesign

Identification of Functional Requirements

10/01/2009

Participants

Staff:

Lorenzo Allori, Technology Director Alessio Assonitis, Research Director

Research Fellows:

Elena Brizio Lisa Kaborycha Stefano Dall'Aglio

Documents to refer to:

- a) Current system slides
- b) Current system slides things to change

In these slides our current Microsoft Access-based (MAPDocSourcesV4D) data entry system is analyzed in all its parts (the PowerPoint slides we used during the meeting are shown after this document)

Definitions of terms:

Hereafter, the terms "slides" or "the slides", refers to *Current system slides*The terms "mslides" or "the mslides", instead, refers to *Current system slides - things to change*

We will also refer to DocBase, PeopleBase and VolBase, which are three of the four parts that our current data entry system is divided in.

<u>Steps</u>

- 1. Introduction
- 2. DocBase mask redesign
- 3. PeopleBase mask redesign
- 4. VolBase mask redesign
- 5. MAPDocSourcesV4D Search Grids
- 6. Printing Functionality

1. Introduction

LORENZO ALLORI

The goal of this meeting is to identify which features we will need to remove and/or change in our new web-based system in order to redesign the new interface properly.

To do this we will check each mask shown in **the slides** to identify which features should be removed and which added, being aware that the program we are going to create is not a client-server application, but rather a web-based one.

2. DocBase mask redesign

2.1. DocBase Extract/Synopsis Tab (slide 1, mslide 1): ALESSIO ASSONITIS

Do you find *Insert or Part No.* (second text box from left) useful?

FELLOWS

It seems that certain inserts have other inserts inside; we should have that text box available even if it is used rarely.

ALESSIO ASSONITIS

Do you find Folio Start No. (second text box from left) useful?

FELLOWS

It is the same as the previous textbox discussed before. It is rare, but we use it.

ALESSIO ASSONITIS

What about the Reckoning checkbox?

FELLOWS

It is used when a dating style is uncertain. We should change its label to Reckoning unsure.

2.2. DocBase People tab (slide 2, mslide 2): FELLOWS

A search button for the portrait indexed in a document should be available inside the PeopleBase.

LORENZO ALLORI

We will discuss this issue during the analysis of the PeopleBase slides.

2.3. DocBase Volume Info tab (slide 3, mslide 3): ALESSIO ASSONITIS

This section should be removed; it should be kept, instead, in the Volbase.

2.4. DocBase Topic tab (slide 4):

FELLOWS

Would it be possible to create a new Topic Category and add it to the existing ones?

ALESSIO ASSONITIS

No. One would have to re-read and re-vet all the volumes that have been entered so far. However, what we can do---and should do---is create a Topic Category, which derives from two or more existing ones. For example, all documents dealing with dance or ballet are entered either under the Topic Categories *Music and Musical Instruments* and/or *Theater and Festival*. In Early Modern Europe, there was no dance performance without music, and usually every theater performance included both dance and music. It is for this reason that Fellows could read over all documents in the categories *Music and Musical Instruments* and *Theater and Festival* and select those documents in which dance is referenced directly and link them to a new Topic Category entitled *Dance and Ballet*.

LORENZO ALLORI

Ok, we will create a new Topic Category called *Dance and Ballet*. It will be added to the topic list.

2.5 DocBase "Fact Check" tab (slide 5, mslide 4):

ALESSIO ASSONITIS

This should be redesigned; we should take out Dates, Places and Person textboxes.

FELLOWS

It should be possible to indicate that a *Fact Check* is available for that document and it should be visible to the community, so that the community users can answer the questions or the issues that are listed in the *Fact Check* section.

LORENZO ALLORI

We can highlight the Fact Check label to identify the presence of a Fact Check related to that particular document.

3. PeopleBase Mask redesign

3.1 PeopleBase Vital Stats tab (slide 6, mslide 5):

ALESSIO ASSONITIS

We should remove the checkbox named "PORTRAIT INDEXED in the DOCBASE".

FELLOWS

We should be able to add images for the portraits. We could use the National Portrait Gallery web site to do this and we should also permit the members of the community to do that.

LORENZO ALLORI

We can also authorize the community users to upload a picture from the National Portrait Gallery. This upload has to be vetted by a Fellow or a Staff member prior to being published online.

3.2 PeopleBase Names tab (slide 7):

ALESSIO ASSONITIS

Ok, nothing to change and/or add.

3.3. PeopleBase Parents tab (slide 8):

ALESSIO ASSONITIS

Ok, nothing to change and/or add.

3.4. PeopleBase Children Tab (slide 9):

ALESSIO ASSONITIS

Ok, nothing to change and/or add.

3.5. PeopleBase Spouses Tab (slide 10):

ALESSIO ASSONITIS

Ok, nothing to change and/or add.

3.6 PeopleBase Roles/Occs Tab (slide 11):

ALESSIO ASSONITIS

Ok, nothing to change and/or add.

3.7. PeopleBase Docs Tab (slide 12):

ALESSIO ASSONITIS

Ok, nothing to change and/or add.

3.8. PeopleBase Biblio Stats Tab (slide 13, mslide 6):

ALESSIO ASSONITIS

This should be restructured. The textboxes should be completely free without having to add a master biblio record.

FELLOWS

We should also remove the button "Add New Master Biblio Record"

4. VolBase mask redesign

4.1. VolBase General tab (slide 14):

ALESSIO ASSONITIS

The textbox called *Internal Memo* should be changed to *Volume Fact Check*.

4.2. VolBase Description tab (slide 15):

ALESSIO ASSONITIS

Ok, nothing to change and/or add.

4.3. VolBase Correspondents tab (slide 16):

ALESSIO ASSONITIS

The *From*: and the *To*: textboxes should be converted into two automatic lists that the program creates itself.

FELLOWS

The names should be clickable so that if you click on a name you get the *PeopleBase* record of that sender or recipient.

In the *Context* tab, one should be able to add a link to a place indexed in the database.

4.4. VolBase Documents tab (slide 17):

ALESSIO ASSONITIS

Ok, nothing to change and/or add.

5. MAPDocSourcesV4D Search Grids

(slides 18-19)

LORENZO ALLORI

As this will be a web-based application the button grids are no longer needed. Web-based programs have a completely different approach: for doing a search over our database we will be creating a search page with all the fields, then we will be able to concatenate those search strings to get the best results.

6. Printing functionality

(mslide 7)

FELLOWS

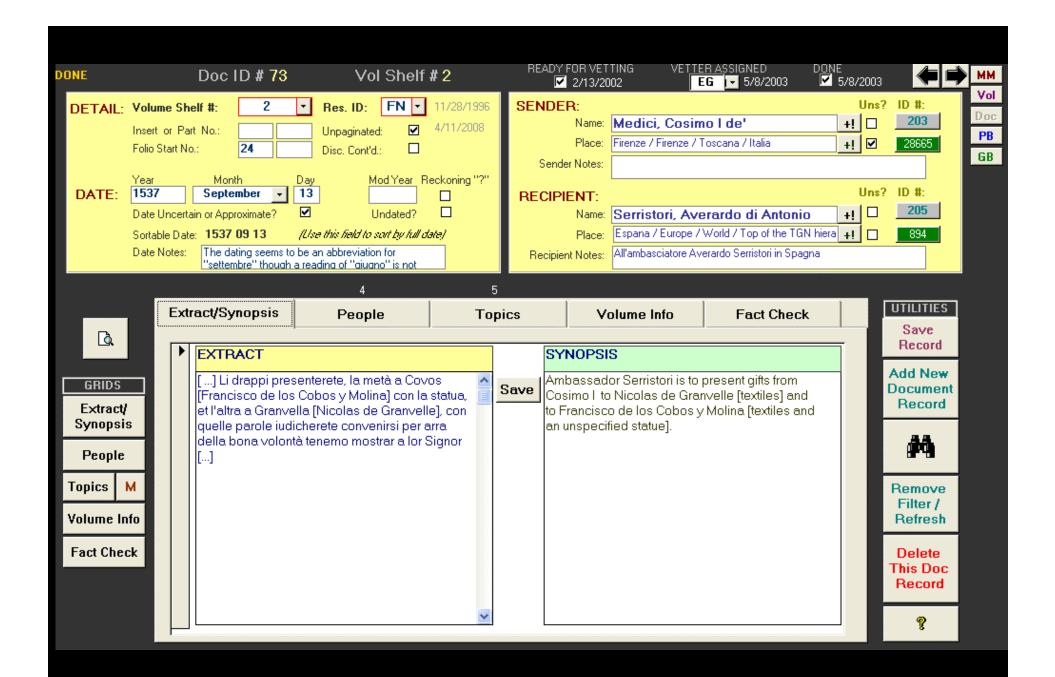
The button for printing documents is very difficult to use and it is a waste of time.

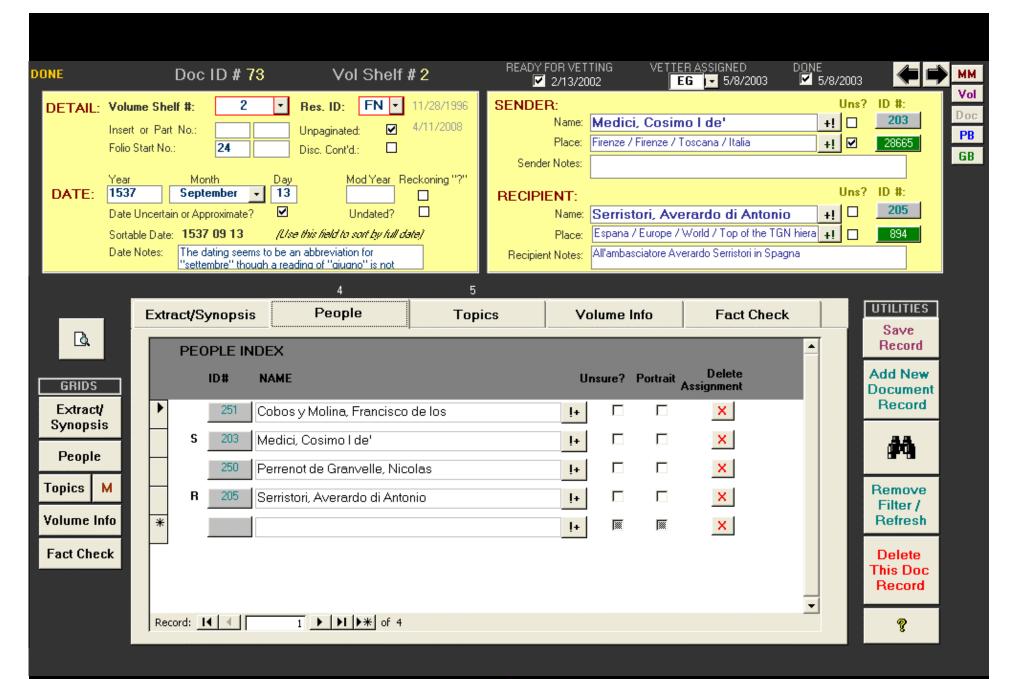
LORENZO ALLORI

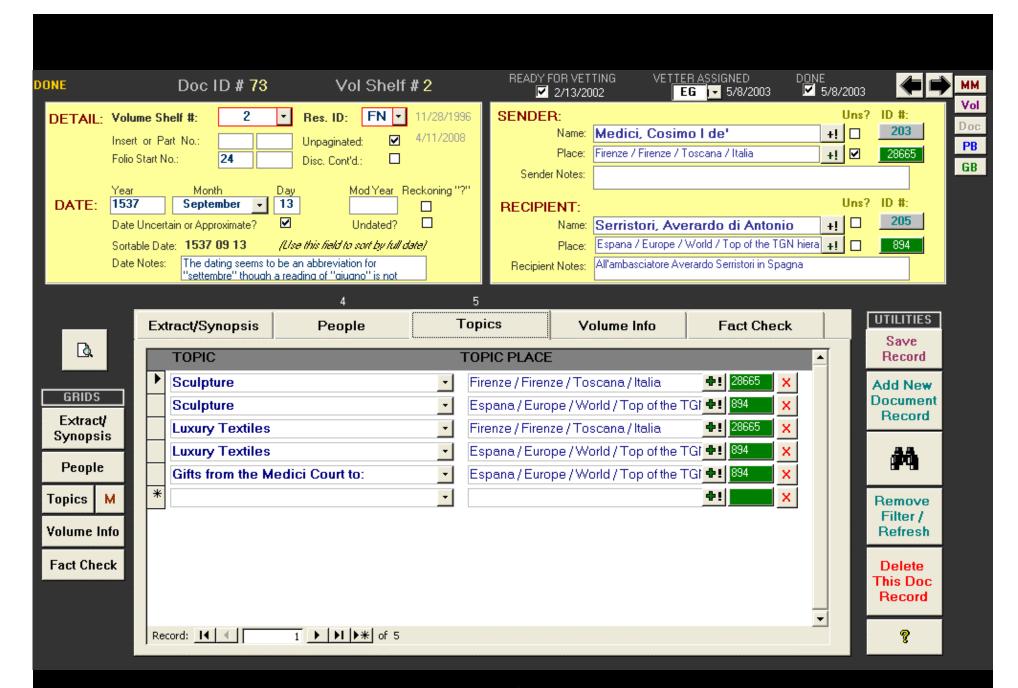
Yes, the printing feature will be implemented in a completely different way. You will have a button for printing the document, volume, bio or place page and also, a printing feature will be implemented on the search page, making it able to print only the selected documents in a reliable way.

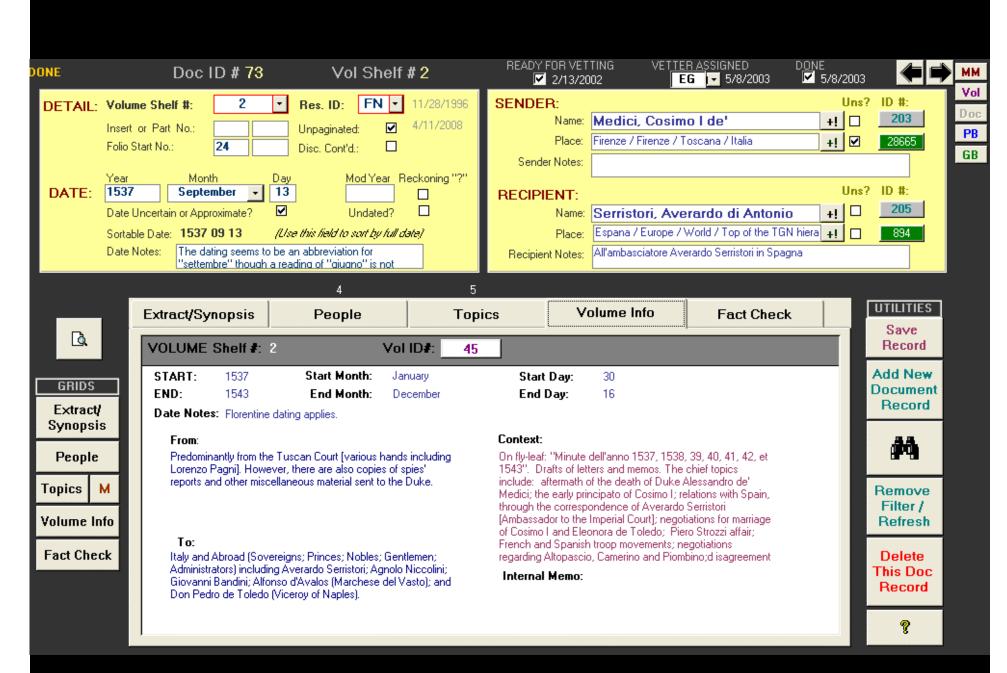
Minutes of Meeting 1

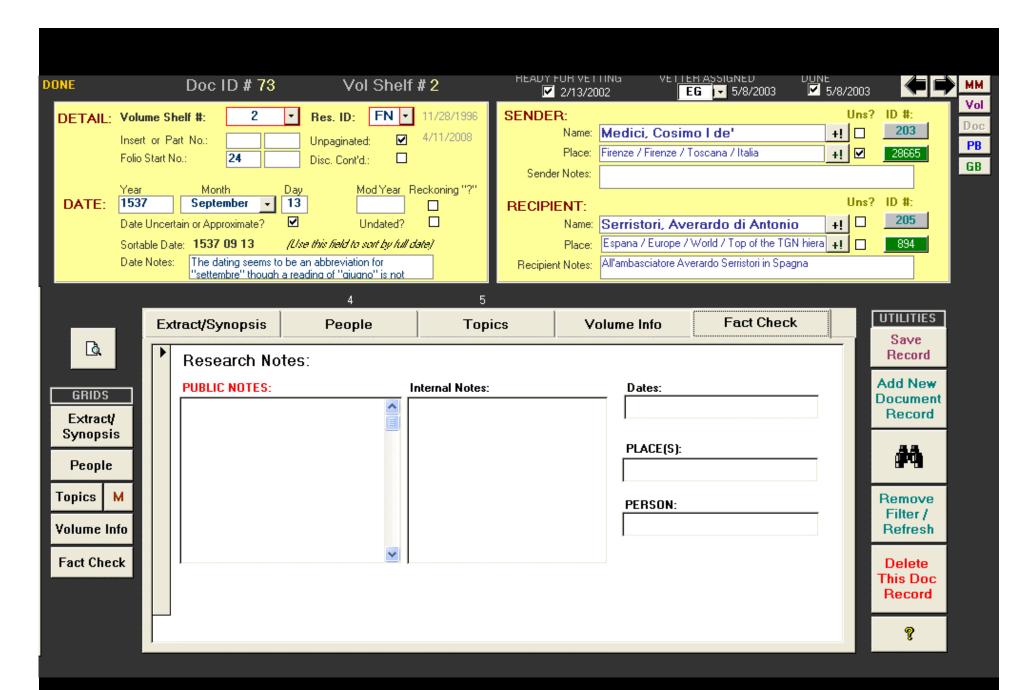
Current System Slides

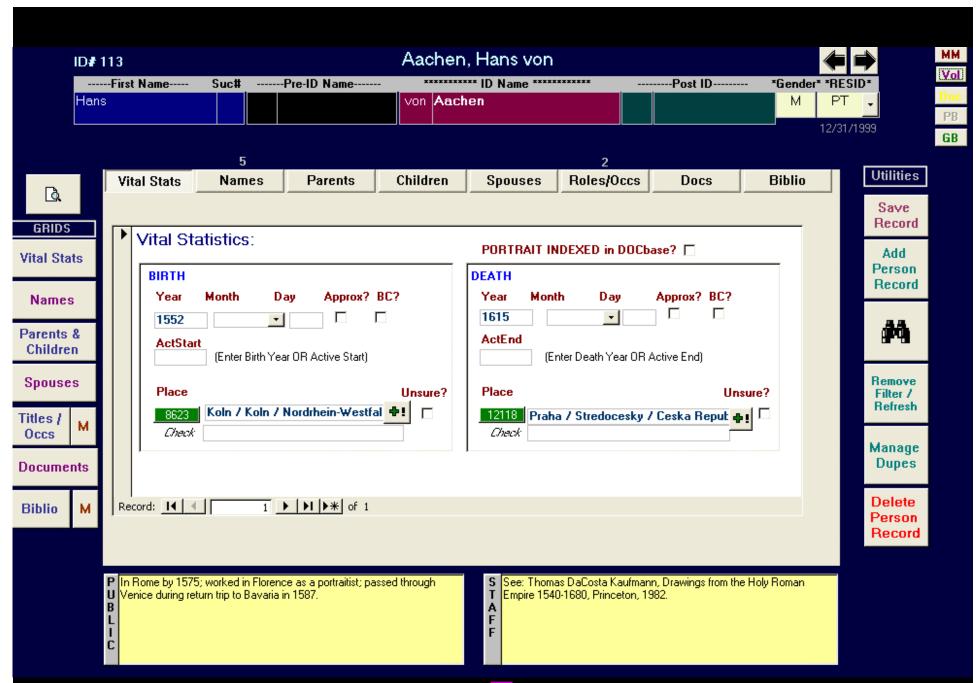


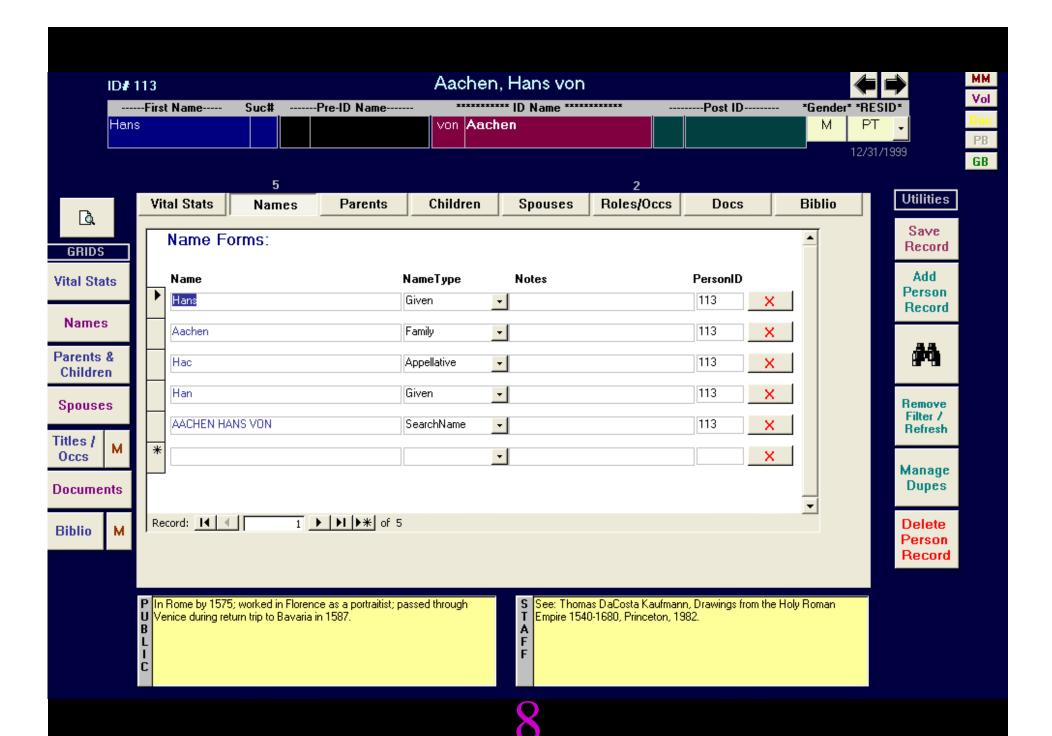


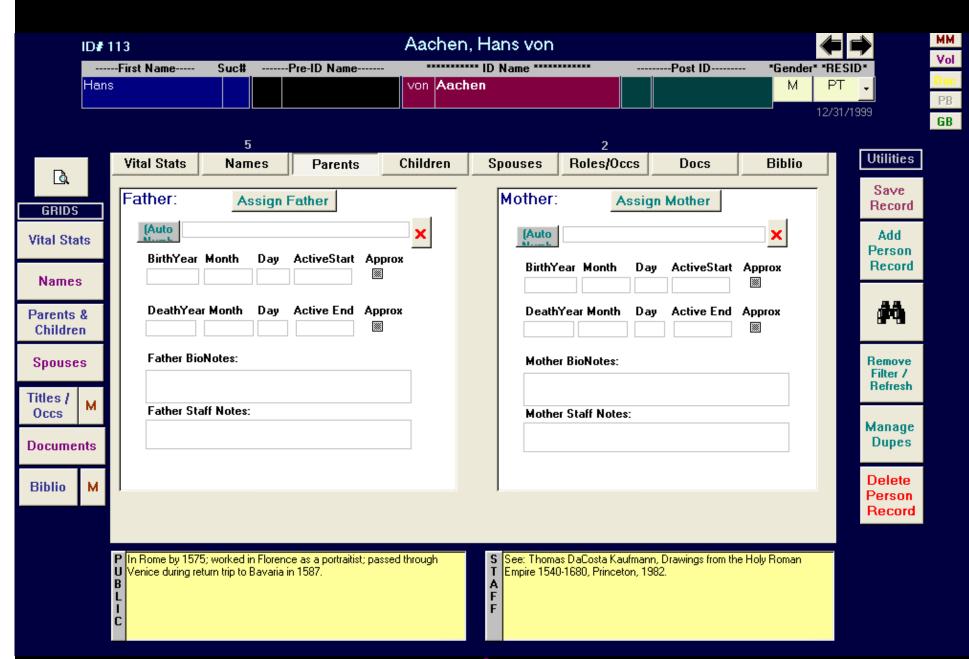


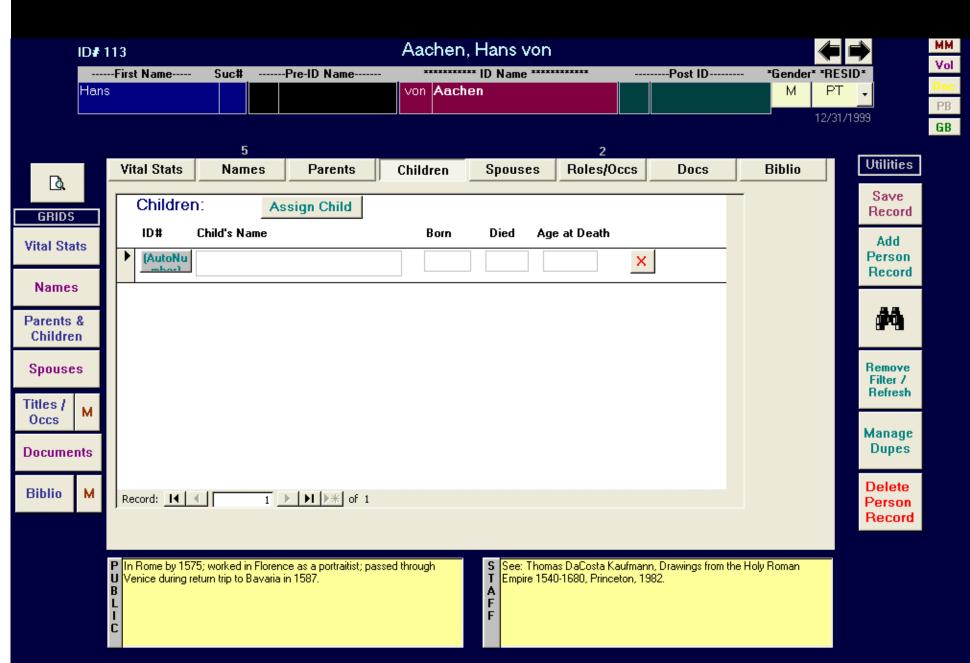


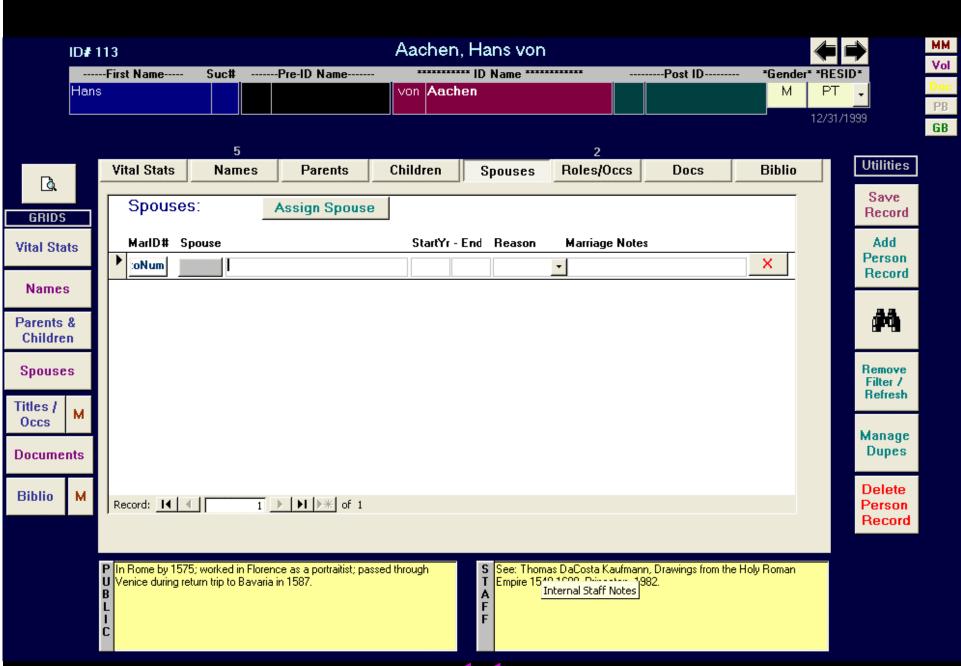


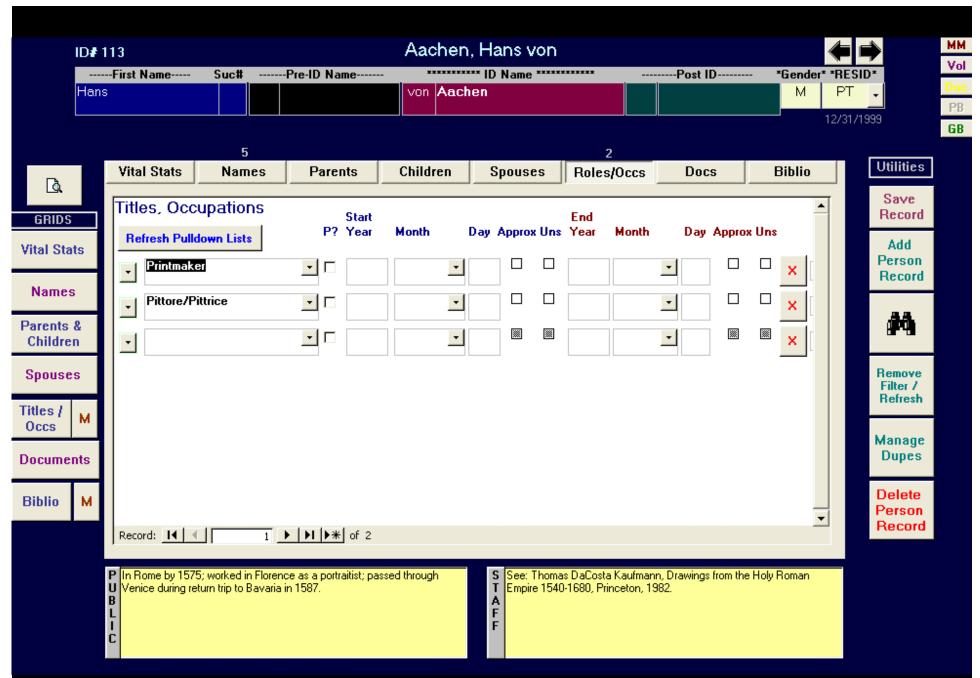


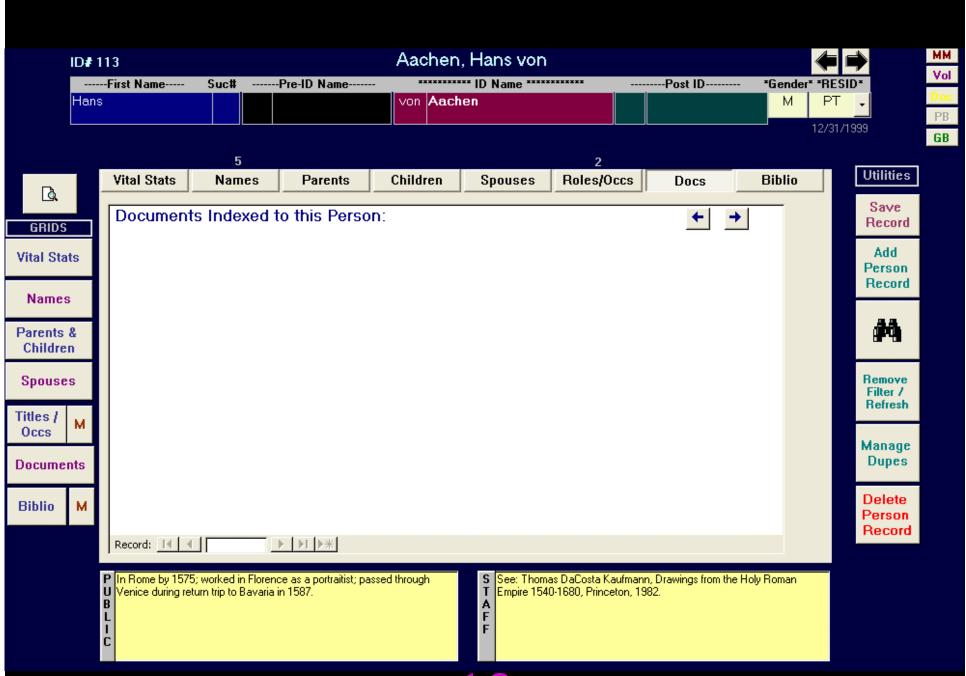


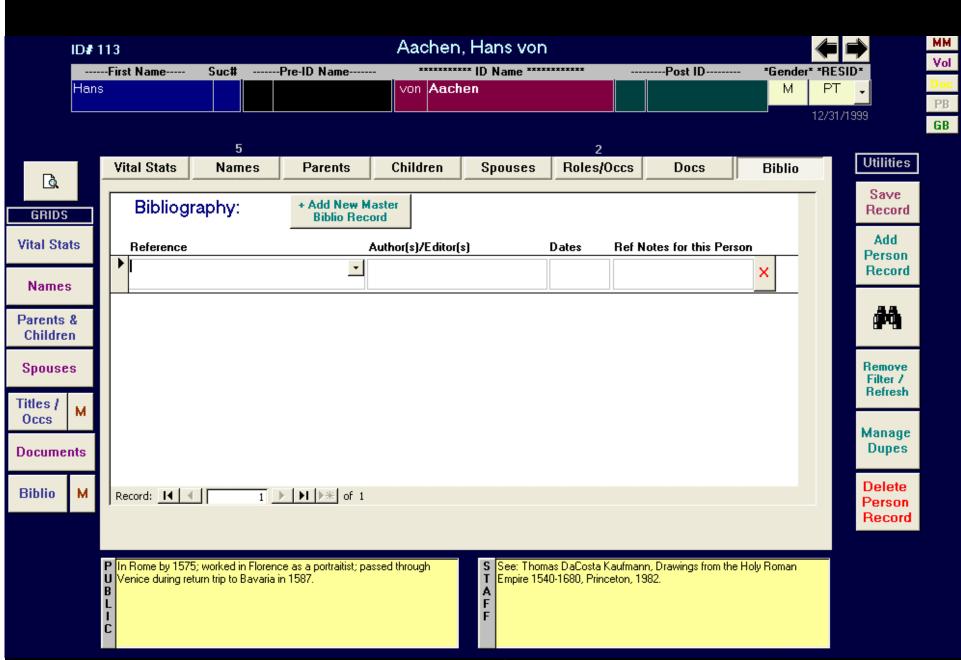


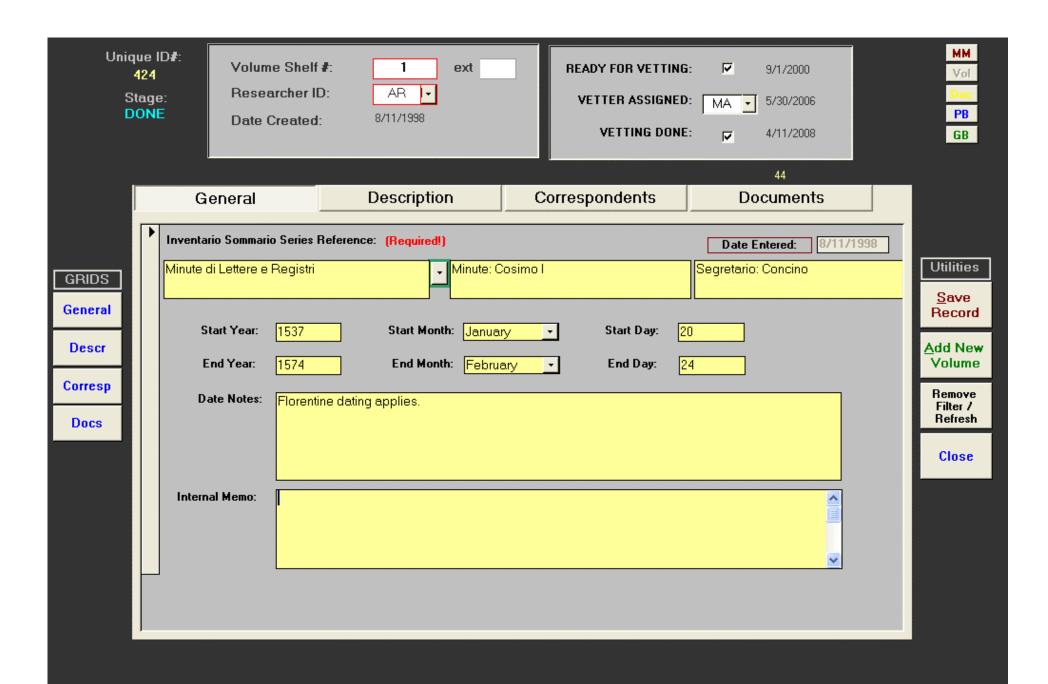


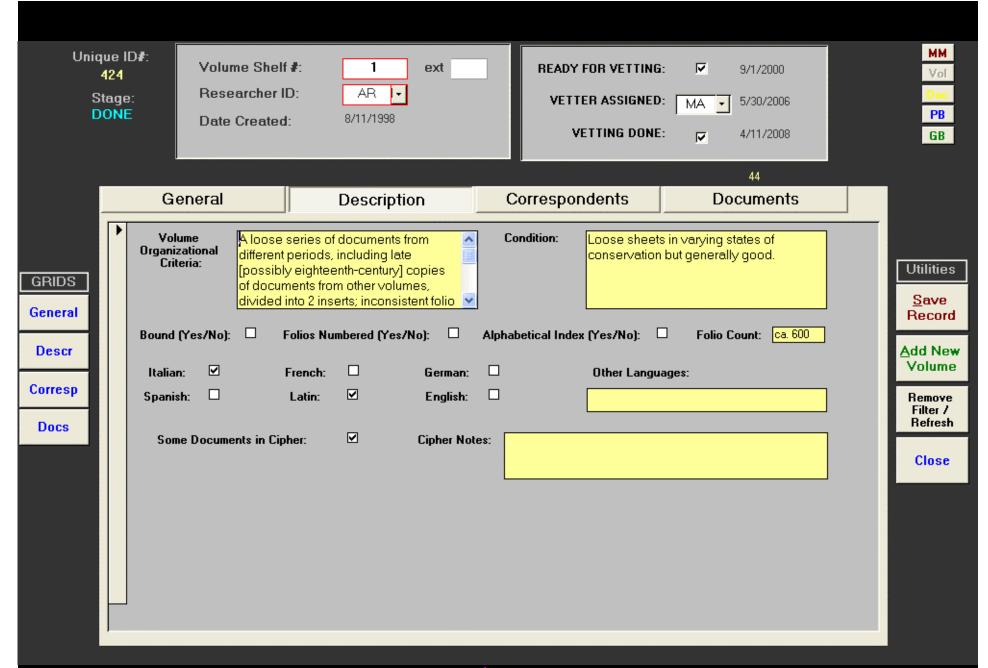


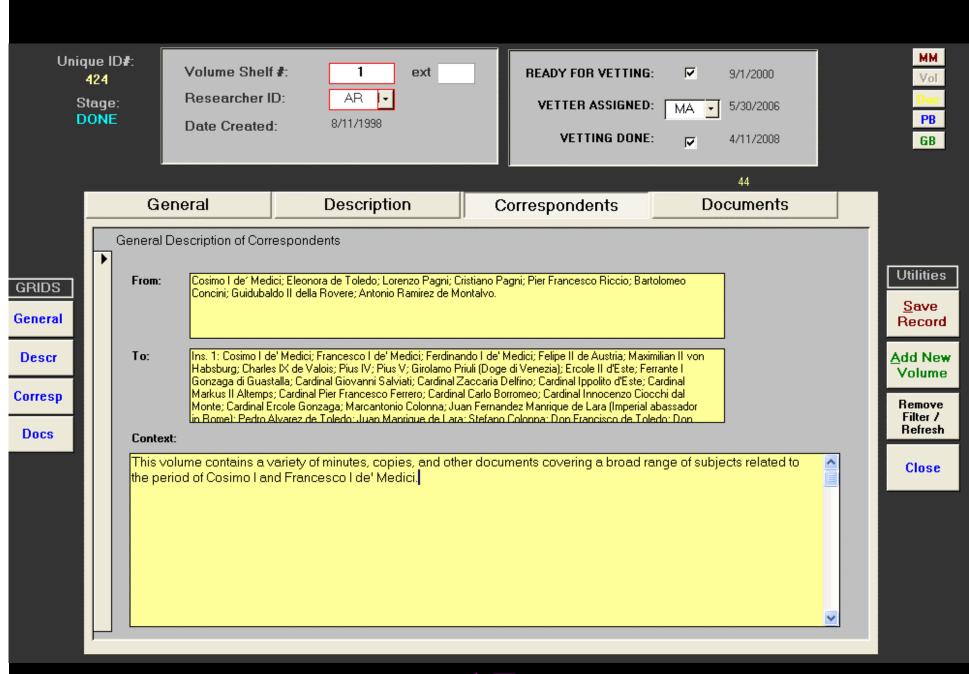




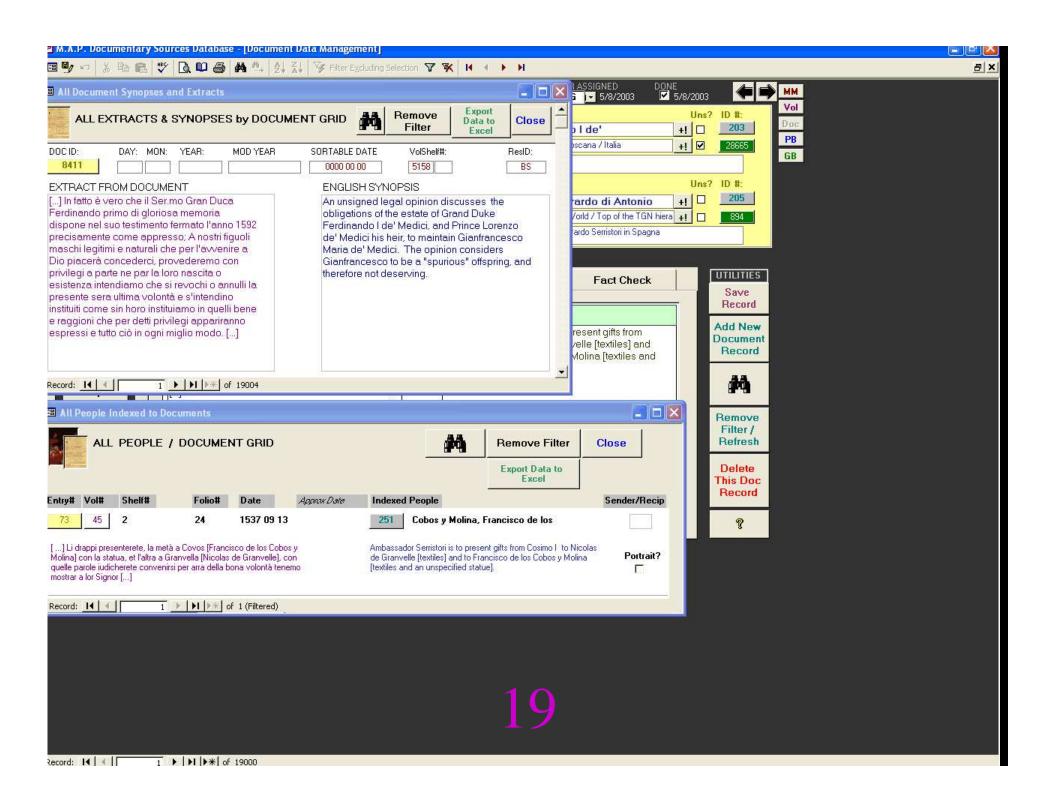


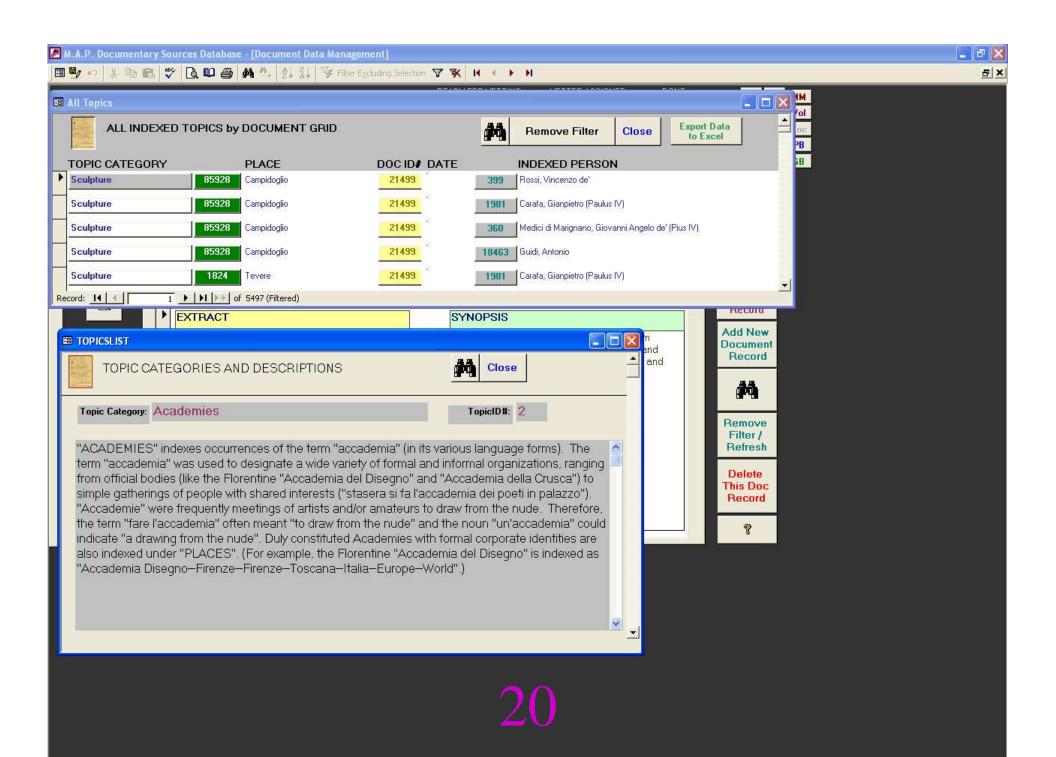






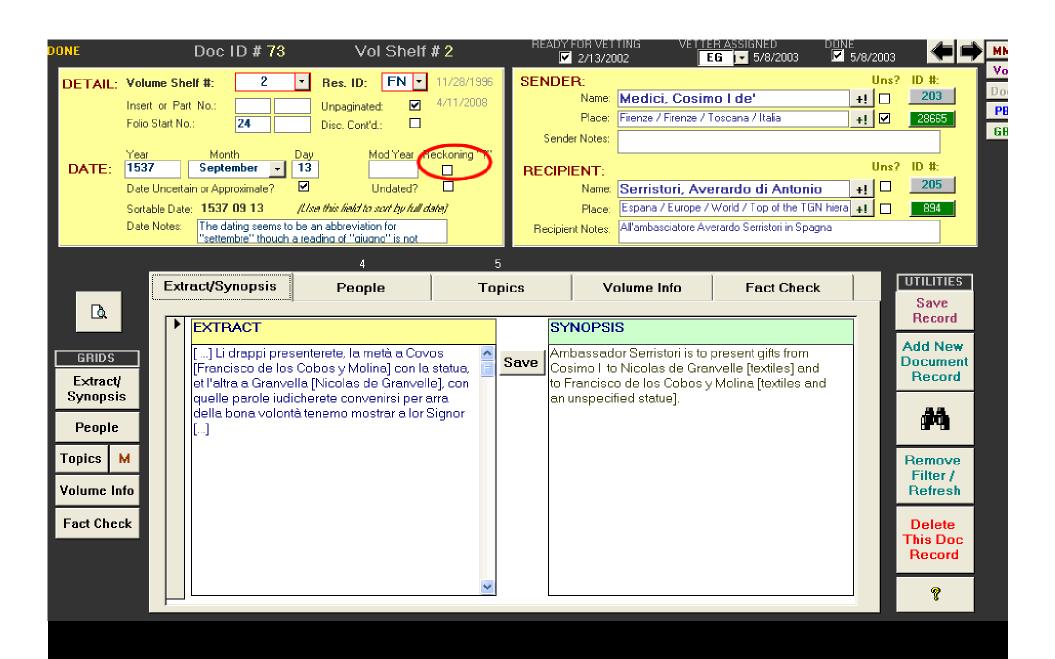


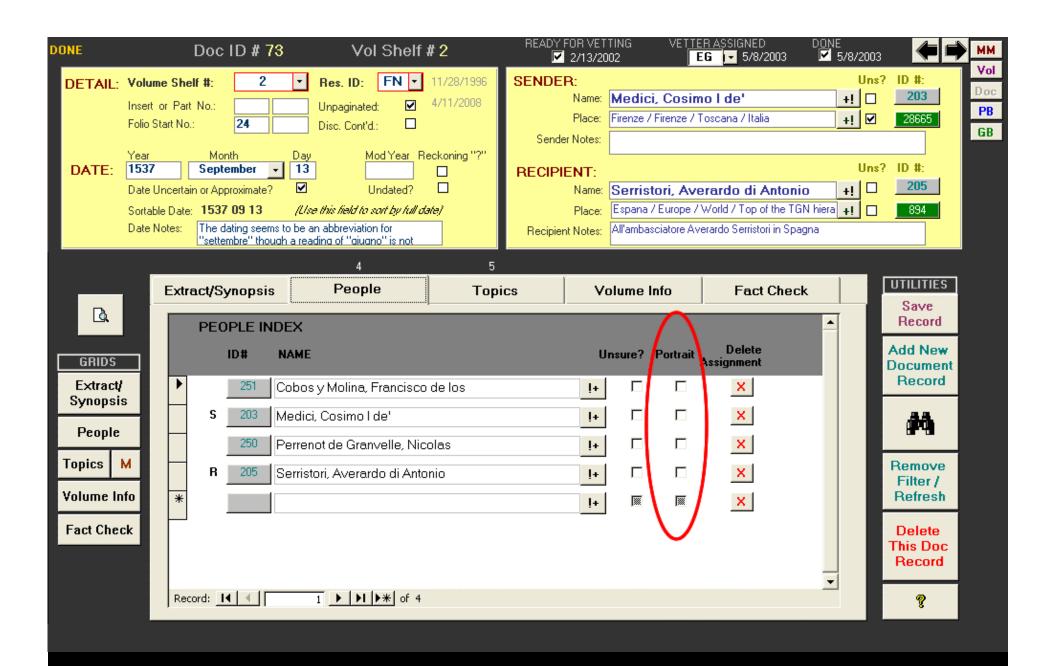


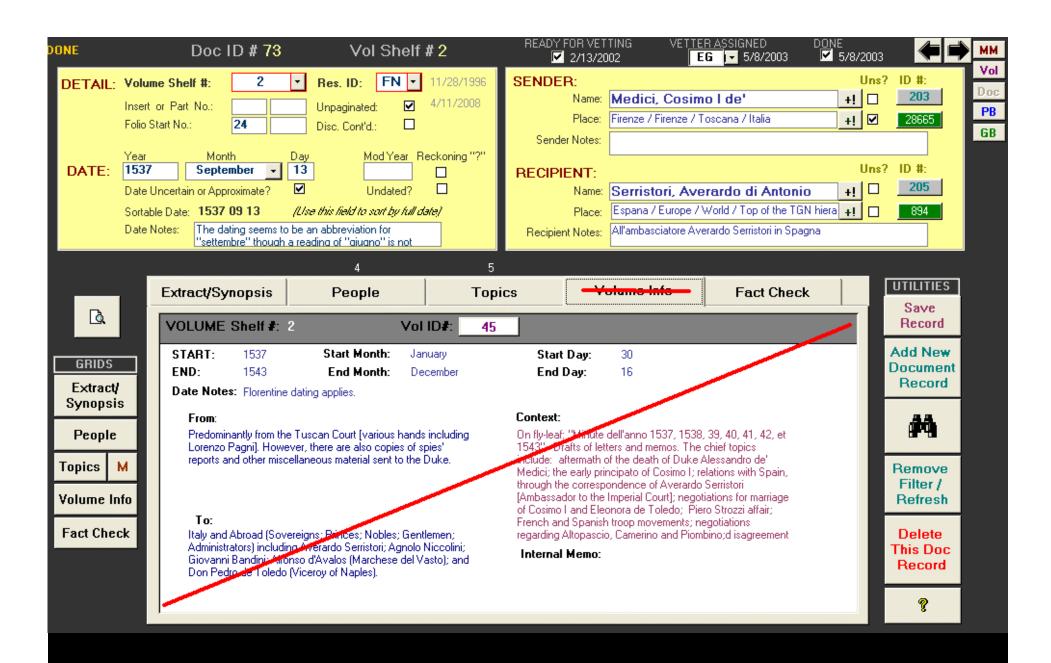


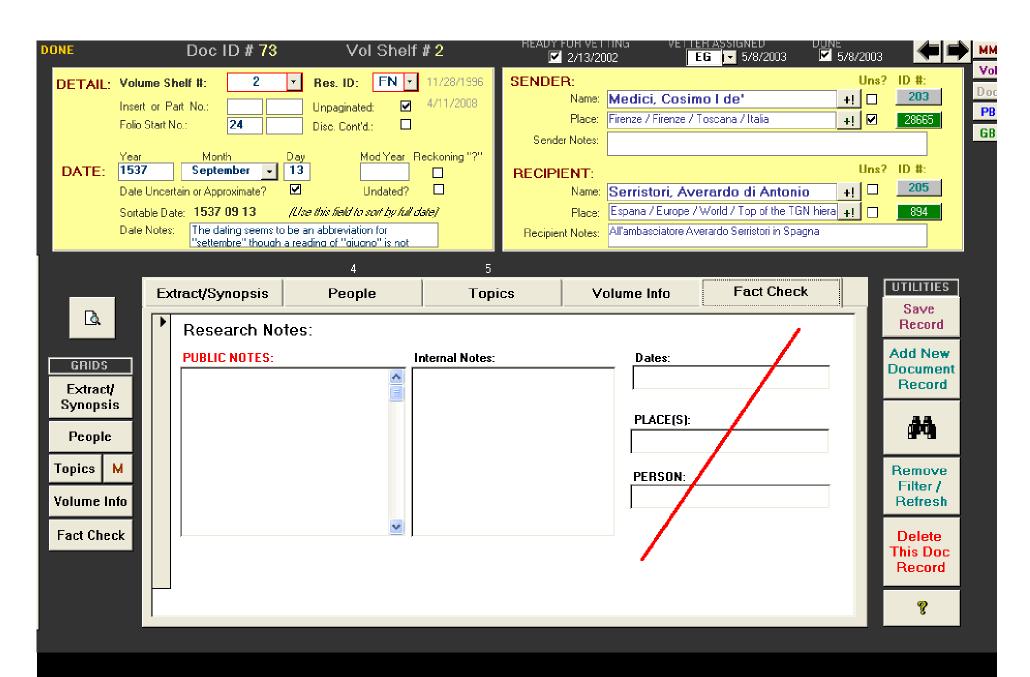
Minutes of Meeting 1

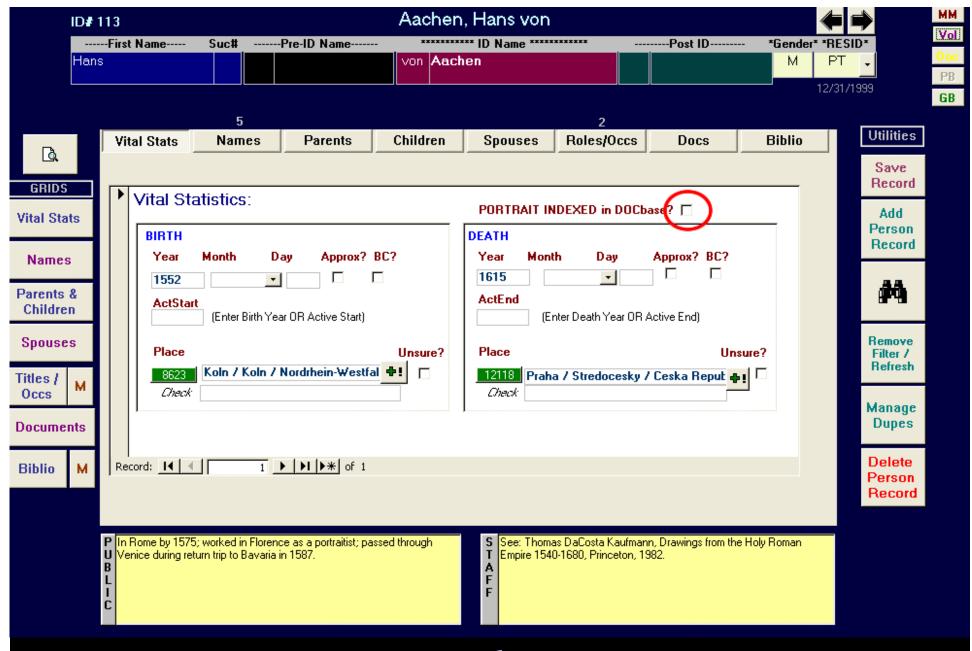
Current System Slides
Things to Change

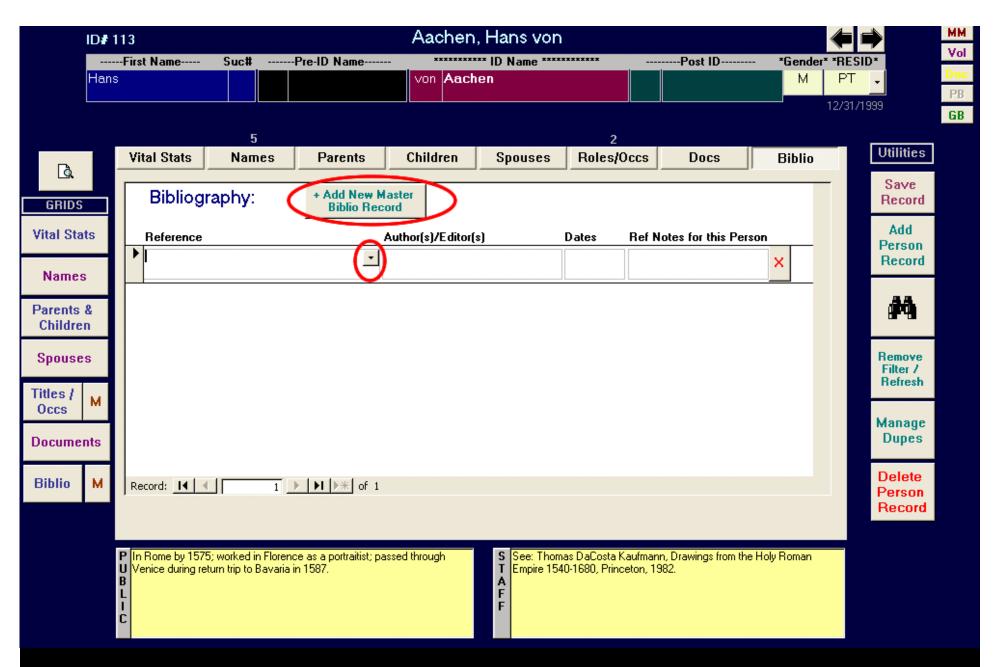


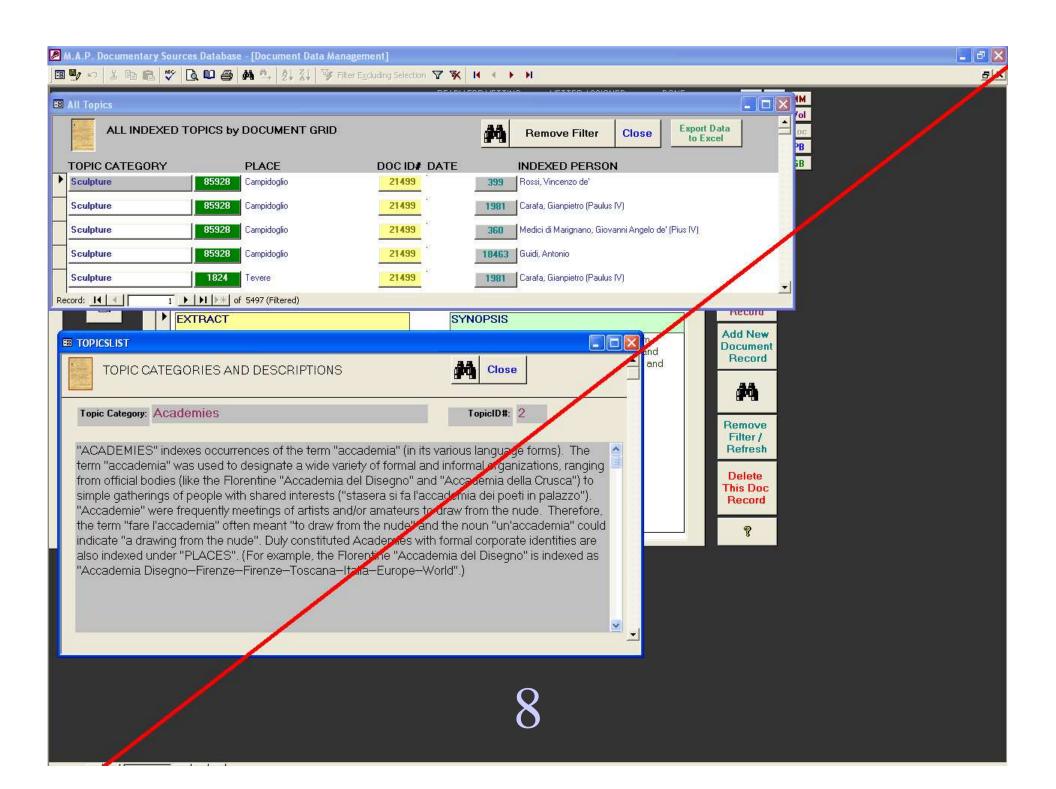


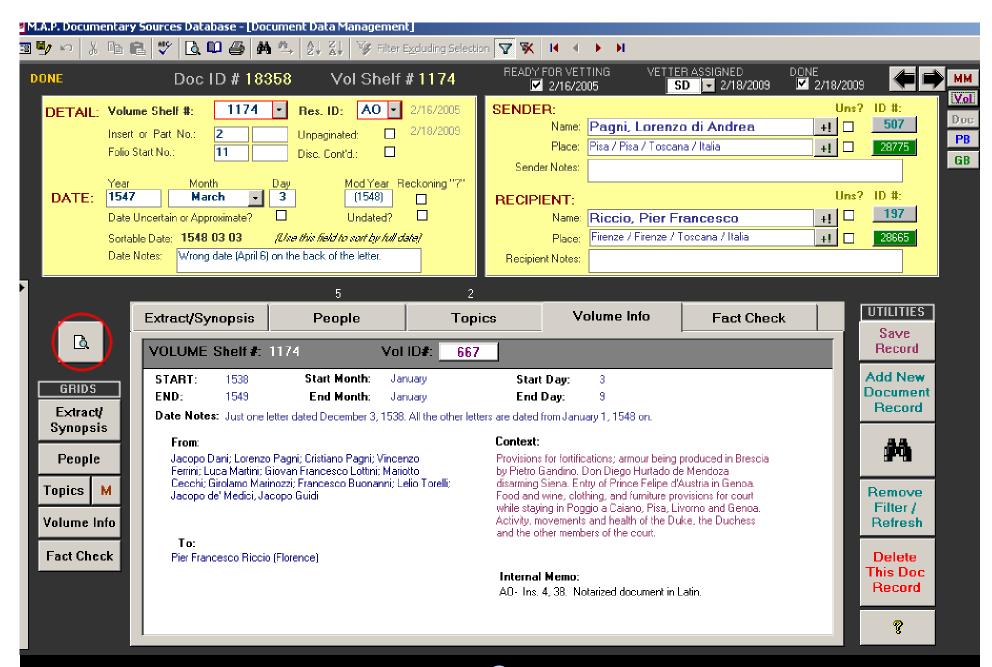












Minutes of Meeting 2 MAP Software Redesign

Digitization process requirements

10/13/2009

Participants

Staff:

Lorenzo Allori, Technology Director Alessio Assonitis, Research Director

Steps

- 1. Introduction
- 2. Document digitization process analysis
- 3. Study of the requirements for the Manuscript Viewer

1. Introduction

LORENZO ALLORI

The goal of this meeting is to:

- Identify which volumes (filze) will be digitized, who will be involved in the process and, how the work will be accomplished.
- Identify the features needed in the web-based program for viewing the digitized manuscripts (Manuscript Viewer).

2. Analysis of the document digitization process

2.1. Already digitized volumes and documents

LORENZO ALLORI

We have already some digitized documents. How many of them do we have?

ALESSIO ASSONITIS

The first 20 volumes are already digitized. These digitized documents will be uploaded to the new system.

2.2. Documents in our current data entry interface

LORENZO ALLORI

How will we deal with the documents in our current database that need their digitized document to be attached?

ALESSIO ASSONITIS

The volumes will be digitized in their entirety starting from the volume number twenty-one. Once a volume will be completed a Fellows will attach the correct manuscript folios to the correspondent indexed document in the database.

2.3. Human Resources

LORENZO ALLORI

We will need a person for the digitization and the upload of documents to our system. The manuscripts need to be handled with care when digitized. We should plan a meeting with the State Archive in Florence to discuss this issue.

3. Study on the requirements for the Manuscript Viewer

LORENZO ALLORI

What features will we need in the manuscript viewer?

ALESSIO ASSONITIS

The digitized documents will have to be readable and therefore should be in high-resolution. Additionally, a zoom-in/zoom-out feature must be available. We should ask the State Archive in Florence if it is possible to use the *manuscript viewer applet* they use in their web - site.

LORENZO ALLORI

Yes, it does also comply with the State Archive in Florence regulations for publishing online digitized manuscripts. During Phase 2, we will find a way to adapt it to our new system.

Minutes of Meeting 3 MAP Software Redesign

Meeting with the State Archive in Florence Officials

10/28/2009

Participants

Staff:

Lorenzo Allori, Technology Director
Marcello Botrugno, Software Architect
Francesca Klein, State Archive in Florence, Director of the Reading Room
Francesco Martelli, State Archive in Florence, Curator of the Medici Granducal
Archive Collection

Steps

- 1. Introduction
- 2. Analysis of the Digitalization Technician workflow
- 3. Definition of the standards for cataloguing digitalized document images
- 4. Analysis of the Manuscript Viewer program

Documents to refer to:

- *Minutes of Meeting 3 - catalog description*, that is shown after this document.

1. Introduction to the document digitalization process

LORENZO ALLORI

The goal of this meeting is to:

- Describe the Digitalization Technician workflow
- Identify and define the State Archive in Florence standards for cataloguing the digitalized documents.
- Analyze the tool for viewing the digitalized document in use by the Florence State Archive to see if it matches our requirements.

2. Analysis of the Digitalization Technician workflow

LORENZO ALLORI

Dr. Klein, would you please describe the workflow of the technicians that you use for digitalizing the manuscripts?

FRANCESCA KLEIN

First the volume must be "prepared": the volume's conditions must be analyzed, identifying if the volume must be restored; a catalog description (see attached

MM3_catalog_description.pdf for reference) of the volume should be compiled and eventually the pages should be numbered in order to link the image to the document. During this phase a member of the State Archive in Florence staff, expert in the Medici Granducal Archive Collection, must assist the Digitalization Technician.

The second step is the actual digitalization. The digitalization technician will be taking color pictures of the document using a camera. The quality should be at 300 DPI minimum. The images should be saved both in TIFF and JPEG format.

The third step regards the image's storing. The images will be downloaded from the digital camera on to a computer, where they will be saved matching our cataloguing standards.

3. Definition of the standards for cataloguing digitalized document images

MARCELLO BOTRUGNO

What are your cataloguing standards for digitalized images?

FRANCESCA KLEIN

Before downloading the images from the digital camera, a directory will be created on the computer, and named as the number of the volume. Then, the images are downloaded to that directory. The format of the names of each image must have: at least four progessive numbers_"coperta/C", G ("guardia")/A(allegato)_three numbers (number of folio)_R/V(recto or verso).

Example: 0005_C_009_R

FRANCESCO MARTELLI

Would it be possible to include the image of the spine of the volume in the MAP new system?

LORENZO ALLORI

Sure, it could be possible to show it in the *VolBase* so that, a volume might also be identified by the picture of its spine. I will discuss this as well with our Research Director, Alessio Assonitis.

4. Analysis of the Manuscript Viewer program

LORENZO ALLORI

Dr. Klein, is it possible to get the sources of the manuscript viewer you are using, as it is an opensource product and it complies with your regulations?

FRANCESCA KLEIN

Yes, I will introduce you to Dr. Parrini from the Scuola Normale di Pisa, who developed our system to view the manuscripts online; you will need to ask him for the sources.

Catalog description

Istituto*	Archivio di Stato, Firenze
Fondo*	
Serie*	
N unità*	
Titolo*	
Date estreme*	
Descrizione contenuto*	
Legatura*	
Supporto*	
Cartulazione*	
Note alla cartulazione*	
Carte bianche*	
Carte mancanti*	
Dimensioni: base (mm)*	
Dimensioni: altezza (mm)*	
Sistema di scansione (marca e modello dello	
scanner, software utilizzato durante il processo	
di scansione)	
Profondità del colore (es: RGB 24 bit)	
Risoluzione	
Formato	
Compressione	
Nome files	ES.:
	quattro cifre (numero progressivo delle
	immagini) _ coperta/C (carta)/G(guardia)/A
	(allegato), _ tre cifre (numero della pagina o
	carta), _ R/V
Numero totale immagini	
Dimensione media immagini	
Dimensione totale immagini	
Responsabile fotoriproduzione	
Data ripresa	
Operatore	

Minutes of Meeting 4 MAP Software Redesign

Phase 1.3 Community requirements

10/29/2009

Participants

Staff:

Lorenzo Allori, Technology Director Alessio Assonitis, Research Director

Steps

- 1. Introduction
- 2. Study of the requirements for the Community module

1. Introduction

LORENZO ALLORI

The goal of this meeting is to identify the requirements for the Community module. This part of the system will allow the Community users to interact with the Fellows' work and amongst one another.

2. Study on the requirements for the Community module

LORENZO ALLORI

Which are the tasks a Community user can perform once logged on?

ALESSIO ASSONITIS

Community users will be able to:

- See read-only data such as: documents including digitalized images, bio records, volume records and place records.
- Comment on each record
- Edit their personal profile
- Communicate through personal messages between each other, or send messages to a MAP Fellow.

LORENZO ALLORI

Comments posted by users do not have to be immediately put online. It is important that a Fellow or a Staff member, who will have the role of checking user comments, monitors each comment posted by a user, and then authorizes the comment or deletes it.

ALESSIO ASSONITIS

The Fellows, who have this role, should be notified on each comment posted.

LORENZO ALLORI

They will be notified via email or through the personal-messaging section the system will provide. This matter will be studied during Phase 2.

Minutes of Meeting 5 MAP Software Redesign

Phase 1.3 History Log module requirements

03/11/2009

Participants

Staff:

Lorenzo Allori, Technology Director Marcello Botrugno, Software Architect

Steps

- 1. Introduction to the History Log module
- 2. Study of the requirements for the History Log module
- 3. Study of the requirements for the History Log module reporting interface

1. Introduction to the history log module

LORENZO ALLORI

The goal of this meeting is to identify the requirements for the *History Log module*. This module will be used to track each user's work as well as the documents' revisions.

2. Study on the requirements for the history log module

2.1. What to log

LORENZO ALLORI

We have to define what we should record in detail.

On to per-user-logging, it is important to collect the Ip addresses and users' "actions".

MARCELLO BOTRUGNO

What do you mean by "actions"?

LORENZO ALLORI

"User actions" refers to the tasks a user can perform, for example, new record creation, deletion of a record, and modification of a particular field.

MARCELLO BOTRUGNO

What about documents' revisions?

LORENZO ALLORI

My recommendation is a *Wikipedia* style revisions, so that each revision of a "document page", a "people page" or a "places page" would be stored in a separate database to keep track of each user's work. This will also be useful for backup purposes.

MARCELLO BOTRUGNO

What about multiple users editing the same document? If two users are modifying the same records we could potentially have multiple errors and duplicated fields in the database.

LORENZO ALLORI

One solution to this is blocking, *record locking*, a document from being edited by more than a single user at a time.

MARCELLO BOTRUGNO

To record the revisions we will create new tables in the database and store this information there. The new tables will have the same record fields of each of the different sections of the database (DocBase, PeopleBase, Volbase, GeoBase).

2.2. How to read logs LORENZO ALLORI

To read these logs, a web-interface is needed.

MARCELLO BOTRUGNO

My recommendation is that the logs, as well as the revision history, should be visible to the system administrators only.

3. Study on the requirements for the history log module reporting interface

LORENZO ALLORI

Two different interfaces will be needed to read these logs: one displaying per-user logs (ip addresses and actions) and the other for viewing prior revisions of the documents.

MARCELLO BOTRUGNO

These two interfaces will be integrated inside our web-based application and accessible only to the system administrators.

LORENZO ALLORI

The fields in the per-user-logging report should be:

- Software's part used (DocBase, PeopleBase, VolBase, GeoBase or other module)
- Time (when the action has been performed)
- Ip address
- User's Full Name
- Action performed

It will be possible to search combining each of these fields to get detailed logs the Administrators might need.

MARCELLO BOTRUGNO

In the revisions interface, it will be possible to search the revisions by "doc id", "people id", "volume id" and "place id", as well as the deleted records.

LORENZO ALLORI

Additionally, a button for viewing the old versions of record should be present on each document, people, volume, and place page. Only Administrators will be able to use this feature.

APPENDIX B

Top Level Software Design Document

(Last Edited - March 23, 2010) Lorenzo Allori and Marcello Botrugno

1. II	NTRODUCTION	5
1.1 P	Purpose	5
1.2 A	rchitectural Design	5
1.3 R	References	6
1.4 O	Overview of the Document	6
2. Se	OFTWARE MODULES DESIGN	8
2.1 A	authentication Module	0
2.1.1		
2.1.2	0 0	
2.1.3	User Registration Form	12
2.2	History Log Module	13
2.2.1	. 8	
2.2.2	Revisions	14
2.2	Search module	15
2.3		
2.3.1		
2.3.2 2.3.3		
2.5.5		
2.4	Data Entry module	19
2.4.1	DocBase sub-module	21
2.4.2	VolBase sub-module	23
2.4.3	PeopleBase sub-module	25
2.4.4	GeoBase sub-module	27
2.5	Community Module	29
2.5.1	•	
2.5.2		
2.5.3		
2.6	Administration Module	3.1
	User Management	
2.6.1		
2.6.3	1	
	-,	
2.7	Digitization module	36
2.7.1		
2.7.2	The Digitization web interface	37
2.8	Manuscript Viewer	38
2.8.1	•	
2.9	Deployment Diagram	40

3. US	SER INTERFACE DESIGN	41
3.1 Us	sabililty	41
3.2	The Main Menu	42
3.2.1	On-Site Fellows and Distance Fellows menu	
3.2.2	Community User menus	
3.2.3	Administration menus	
3.2.4	Digitization menu and its web interface	
3.3	Split Screen	48
3.4	The Document Transcription	51
3.5	Getting info from dropdown auto completing menus	55
3.6	How to print records	50
3.6.1	Printing a single item page	
3.6.2	Printing document records from a search list	
0.0		
3.7	Compatibility	61
4. D A	ATA STRUCTURAL DESIGN	62
4. D	ATA STRUCTURAL DESIGN	02
4.1 Da	ata Structure	62
4.2	Current Database Tables	63
4.3	New database tables	70
4.3.1	User action tracking	70
4.3.2	Database Items revisions	70
4.3.3	Community Comments	70
4.3.4	Personal Messages	
4.3.5	Showing digitized manuscripts	
4.3.6	Google Maps integration	
4.3.7	Catalog descriptions	
5. H	ELP SYSTEM DESIGN	72
5.1 Th	ne User Manual	72
5.2	The embedded Help System	73
5.3	The Developer Manual	73
6. AF	RCHITECTURAL STRATEGIES	74
6.1 Fr	ameworks	74
6.1.1	Spring MVC	
6.1.2	Security and Authentication	
6.1.3	Data Model	
6.1.4	System Logs	
6.1.5	Graphics	
J.1.J		

6.	.1.6 Google Maps integration	76							
6.2	Database	76							
6.3	Software Project Management and development tools	76							
6.4	Application Server76								
6.5	Bug-tracking	76							
6.6	Server-side scripting	77							
7.	APPLICATION ENVIRONMENTS	78							
7.1	The Development Environment	78							
7.2	The Testing Environment	78							
7.3	The Production Environment	79							
8.	SERVER SYSTEM	80							
8.1	Servers deployment	80							
8.	.1.1 Servlet container: Tomcat configuration								
8.	.1.2 Database: Mysql configuration								
8.	.1.3 Backup and Restore Procedures								
8.	.1.4 Hardware and bandwidth upgrading	82							

1. Introduction

1.1 Purpose

This document contains the complete description of the Medici Archive Project's (hereafter referred to as MAP) Software System Redesign.

The primary audience of this document is the Software Developers.

NB: This document will be used by the Development Team (hereafter referred to as DT) as a guideline, but it is requested that all members of the DT take an active part in the implementation choices. By "active part" we mean that eventual problems that have not been investigated during the planning phase will be analyzed afterwards, along with the Software Architect's contribution.

1.2 Architectural Design

When developing "data-centric" applications, the most common approach is to divide the software system into three different application layers. Each layer performs specific tasks. In turn, it asks the other two layers to perform different tasks.

These three layers are organized hierarchically as follows:

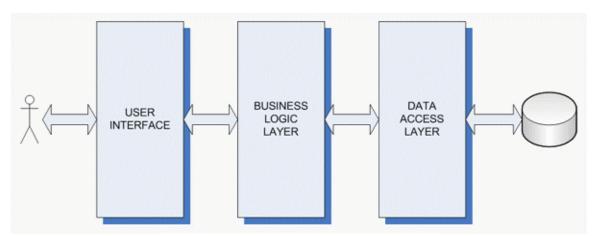


Image 1 – The software application layers

User Interface: this layer handles the system interaction with the users. It includes the readonly and the data-entry masks, the controls, and the mechanisms that appropriately intercept the events that the user performs.

Business Logic Layer: it includes the set of rules that handle information exchange between the Data Access Layer and the User Interface.

Data Access Layer: provides simplified access to the data kept in a persistent storage, such as a RDBMS.

This model will be used to develop the new MAP software system.

1.3 References

Principal Documents:

Appendix A - Top Level System Requirements Document

Minutes of meeting:

Minutes of Meeting 3 – State Archive in Florence requirements

Current websites:

http://documents.medici.org (MAP online read-only database) http://courses.medici.org (MAP online teaching software)

Current Data Structure:

Database iter_medici, our current mysql database tables structure (the data source of http://documents.medici.org) that is shown after this document.

1.4 Overview of the Document

The document is divided into six main sections:

- The *Software Modules Design* section explains how the modules and sub-modules that compose the MAP system are structured and how they will be developed
- The User Interface Design section shows the web interface "look" and its new features
- The *Data Structural Design* section gives a general view of MAP's current database tables as well as the new database tables, which will be created
- The *Help System Design* section shows how the help system works and where the users can access it
- In the Architectural Strategies section the coding approach is explained in detail, as well as the tools that will be used to develop the software platform
- The Application Environments section describes the Development Environment, the Testing Environment, and the Production Environment. This includes the Developers' laptop software configuration as well as the Servers' software

- The *Server System* section describes both the server deployment, as well as the backup and restore procedures

NB:

- The manuscripts and the text displayed on the screen images attached to this document do not correspond to actual documents, but are only samples.
- The images shown in the document are not to be considered the final look of the new web application, but only a suggestion of potential screens and masks, which might be adopted. Additionally, the purpose of these images is to better understand the concepts and the functionalities explained in this document. Referring to these images will facilitate the work of the software developers.

2. Software Modules Design

The web-based application MAP is going to develop has a modular structure:

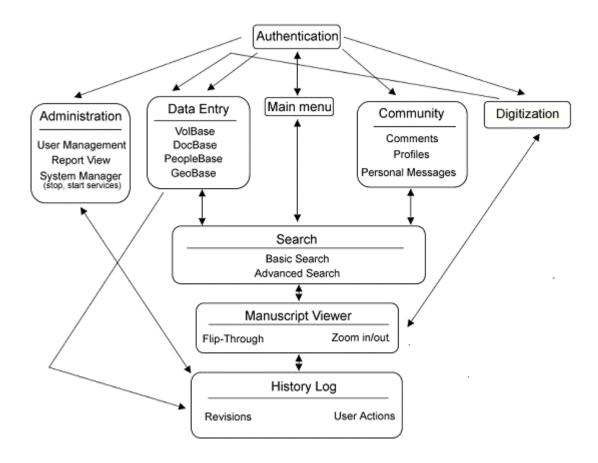


Image 2 – System Modules and Sub-modules

Some modules have sub-modules, such as the Administration module, which is divided into User Management, Report View and System Manager.

The next sections describe each module and its sub-modules in detail.

2.1 Authentication Module

This module recognizes the user and gives him the authorization to use one or more software modules based on the credentials of his users' group.

2.1.1 User Groups

The defined user groups are: Administrators, On-Site Fellows, Senior Distance Fellows, Distance Fellows, Digitization Technicians, Community users and Guests. Some special role users are also present: Distance Fellow Coordinator, Community Coordinator and Digitization Coordinator.

For a detailed description of the each user group see the 2.4 section of *Appendix A - Phase1_TLSRL.pdf - Top Level System Requirements Document.*

2.1.1.1 Administrators

Users belonging to this group are able to use all the Administration Module's sub-modules: *User Management*, Reports and Revisions and System Management [see 2.6].

2.1.1.2 On-Site Fellows

Users belonging to this group are able to use the Data Entry Module [see 2.4] and have the permission to correct and vet other documents entered by On-Site and Distance Fellows. They are part of the Scholarly Community and have also access to all the other modules, with the exception of the Digitization Module [see 2.7] and the Administration Module [see 2.6].

The On-Site Fellows can be assigned (or "flagged") as Distance Fellows Coordinator or Community Coordinator or Digitization Coordinator [see 2.1.1.8].

2.1.1.3 Senior Distance Fellows

Users belonging to this group are experienced scholars who have been trained by the On-Site Fellows on how to use the system. They are able to use the Data Entry Module [see 2.4] and have the permission to correct and vet other documents entered by Distance Fellows. They are part of the scholarly Community and also have access to all the other modules, with the exception of the Digitization Module [see 2.7] and the Administration Module [see 2.6].

The Senior Distance Fellows can be assigned (or "flagged") as Distance Fellows Coordinators and/or Community Coordinators [see 2.1.1.8].

2.1.1.4 Distance Fellows

Users belonging to this group are able to use the Data Entry Module [see 2.4]. Unlike the On-Site Fellows or the Senior Distance Fellows, Distance Fellows are able to enter data but not to vet any document or volume, or delete any content, except for their own entries. They

are part of the scholarly Community and also have access to all the other modules, with the exception of the Digitization Module [see 2.7] and the Administration Module [see 2.6].

2.1.1.5 Community Users

Users belonging to this group can access the Search module [see 2.3]. They can also access the Community module's three sub-modules: comments [see 2.5.1], profiles [see 2.5.2] and personal messages [see 2.5.3].

2.1.1.6 Digitization Technicians

Users belonging to this group are allowed to access the Digitization module [see 2.7] and are in charge of uploading the digitized documents to the system. The Digitization Technician also belongs to the Community User Group and thus has a read-only access to all the items contained in the database and is able to post comments.

2.1.1.7 Guests

These are the users that browse MAP's system without using a username and a password. They can access the Search module, though access to the Manuscript Viewer module [see 2.8] is forbidden. They can, however, fill in the *User Registration Form* [see 2.1.3] to become regular Community Users. Every request is checked by the Community Coordinator and, if accepted, the user becomes a regular Community User [see 2.1.1.5].

	ADM	DE	DEV	DEVD	SRC	MVIEW	СМ	HL	DIG
Adminstrators	\	1	1	1	1	√	1	1	1
On-site Fellows		1		1	1	1	1		
Senior Distance Fellows		1		1	1	√	1		
Distance Fellows		1			1	1	1		
Community Users					1	1	1		
Digitization Technicians					1	1	1		1
Guests					1				

Modules Description:

ADM= Administration module
DE= Data Entry module
DEV= Data Entry module (vet and change or delete any document)
DEVD= Data Entry module (vet and change or delete any Distant Fellows documents)
SRC=Search module
MView=Manuscript Viewer module
CM=Community module
HL=History Log module
DIG=Digitalization module

Image 3 – User Groups summary

2.1.1.8 Special Roles

- The <u>Distance Fellows Coordinator</u> can check the Distance Fellows' work by accessing the *History Log module's* sub-module called *User Actions* [see 2.2.1]. He will be notified in his inbox [see 2.5.3] about any new "ready to be vetted" [see 3.2.1] document created by the Distance Fellows.
- The <u>Community Coordinator</u> can moderate the Community Users' comments. He is notified in his inbox about any new comment posted by the Community Users [see 2.5.3]. He is also notified of any user that fills in the *User Registration Form* [see 2.1.3] and therefore he is in charge of whether or not to allow a user who requests to become a Community User.
- The <u>Digitization Coordinator</u> is notified in his inbox about any new uploaded digitized volume.

2.1.2 The Login Page

On the login page the user can perform three actions:

- Login with username and password
- Login as Guest
- Register to get username and password



Image 4 – Login page

2.1.3 User Registration Form

When a Guest user [see 2.1.1.7] wants to become a Community User, he has to fill in a *User Registration Form.* he can access it from the user login page [see *Image 4*].

Registration Form								
First Name:*								
Last Name:*								
City:								
Country:*								
Email:*								
Re-enter your email:*								
Address:								
Organization:								
Title:								
Interests: *								
Verify that you are a human, WDB fancy captcha ►								
drag pencil into the circle. DROP								
MOT HERE								
Please read this disclaimer and check the 'I agree' box ONLY if you agree to the terms!								
l agree □								
Submit Registration								

Image 5 – User Registration page

This form has the following fields to be filled in:

First Name *
Last Name *
City
Country *
E-mail *
Address
Organization
Title

Interests*

The fields with * are compulsory.

A verification image, like a "captcha", is present in order to protect the forms against spam robots.

A community rules disclaimer is shown, and the user has to read and accept it before submitting the request.

After submitting the request the user receives his username and password by email.

To see how the Authentication module will be developed, see 6.1.2.

2.2 History Log Module

2.2.1 User Actions

This sub-module keeps track of any user action. Once a user logs on, the User Actions sub-module starts logging each action he performs.

Each log entry shows these fields:

Date and time

It address

Username

Full name

Action

Information

2.2.1.1 Date and time

This is a record of the date and time when the log occurs.

2.2.1.2 Full name

This is the full name of the user who performs the logged action.

2.2.1.3 Action

This is the action that a user performs: (logging in, delete, modify, add new record, etc.)

2.2.1.4 Information

This field concerns where the user performed the action: the module or the sub-module that the user was involved in. If modifying, deleting or creating a record, the record ID is shown. If the log concerns the deletion of any volume, bio, document or place, a link to the last revision [see 2.2.2] of that document is available.

For a detailed description of the User Actions database table see 4.3.1.

2.2.2 Revisions

This sub-module records any previous version of each volume, document, bio or place present in the database creating a history of revisions.

Every time that a user modifies or deletes a record, a copy of the previous version of the record is registered in the database [see 4.3.2].

Each document, bio, volume or location page, if accessed by a member of the Administrators user group, has a "history of revisions" link, which can be clicked to see that item's list of revisions.

A tool to browse the revisions can be used by Administrators group users, through the Administration interface [see 3.2.3].

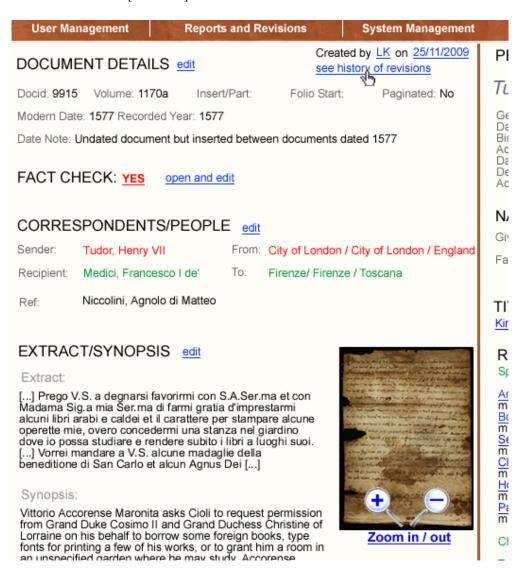


Image 6 – see item's history of revisions

2.3 Search module

This module is used to search through the database records. Everyone can use it and perform three main actions: perform a search, browse through the results, and show the desired items.

2.3.1 Basic Search

The basic search is performed through the Main Menu [see 3.2.1] and permits "a word search" on volumes, documents, people, locations, and Community comments [see 2.5.1]. The results are displayed in a list [see *Image 7c*]. It is also possible to search through the results obtained from a basic search, but only using the *Advanced Search* format.

2.3.2 Advanced Search

The Advanced Search tool permits a fully customizable and accurate search to be performed on the four different items present in the database. Each item has different attributes, thus four different search masks are present: Search Documents, Search Volumes, Search Locations, and Search People. In addition, an option to search through Community comments is present [see 2.5.1 and *Image 28* at 3.2.1.4]. After choosing what to search, the user can create a set of filters combining and concatenating any type of content.

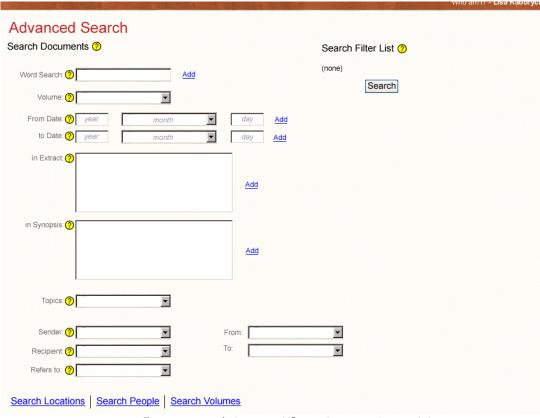


Image 7a – Advanced Search page (sample)

On the left part of your screen you create the filters (by using the "add" buttons); on the right side, you can customize or remove some filters:

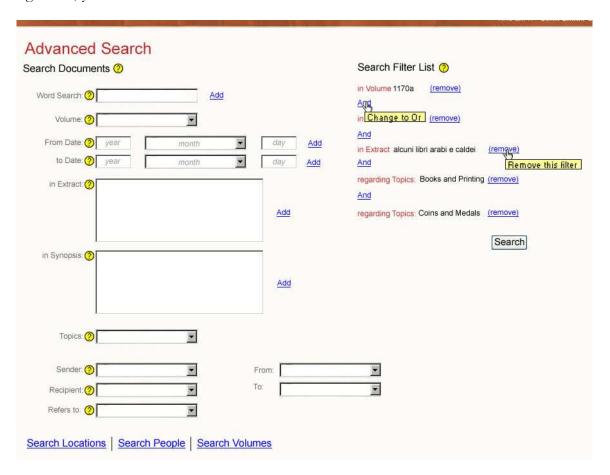


Image 7b - Advanced Search page, filters creation (sample)

The search results will be displayed as follows:

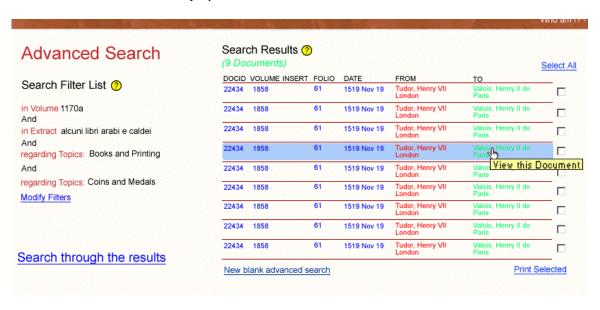


Image 7c – Advanced Search page, search results (sample)

As for the Basic Search, an option for searching through the results is available.

2.3.3 Item file

Clicking on any of the results takes the user to the item file requested. This can be a document, a volume, a bio, or a place. By clicking on the button "turn editing on", the Fellow enters the Data Entry module [see 2.4]. If the user belongs to a group, to whom editing of an item is forbidden, the "turn editing on" button is hidden.



Image 8 – the Turn editing on button

2.4 Data Entry module

This module consists of the data entry interface that On-Site Fellows, Senior Distance Fellows, and Distance Fellows use.

Through this module they are able to create new item records (volume, document, bio or place), modify and/or delete them. Each type of item has its attributes (fields) that can be edited.

The Fellows enter this module by clicking on the button called "turn editing on" (see *Image 8* on the previous page).

Once having entered the Data Entry module, the user can edit each editable part by clicking the "edit" web-link visible at the right side of each section title:

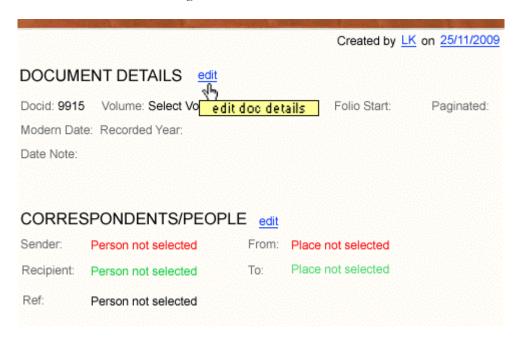


Image 9 – the Edit section web-link

After pressing the edit web-link, the "Edit Box" emerges, allowing the user to edit that section. In the example below, the user is editing the "Document Details".

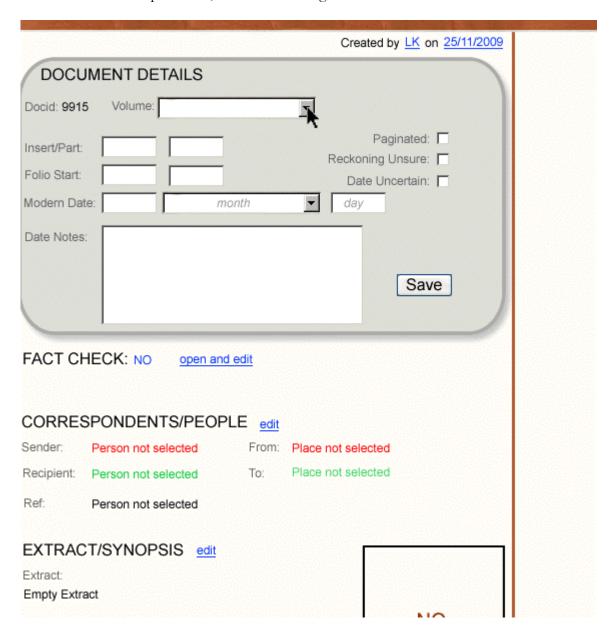


Image 10 – the Edit Box

After editing each section the user presses the Save button and automatically saves and exits the "Edit Box", allowing the user to edit another section.

These operations need to be done without the web page being refreshed, as specified in the requirements documents. Therefore, to create and show the "Edit Boxes" the DWR API along with JQuery will be used [see 6.1.1 and 6.1.5].

2.4.1 DocBase sub-module

This is the data entry part that stores identifying and indexing information regarding the documents themselves.

Using this interface, a scholar is able to transcribe a document record, contextualizing it by linking it to people and locations.

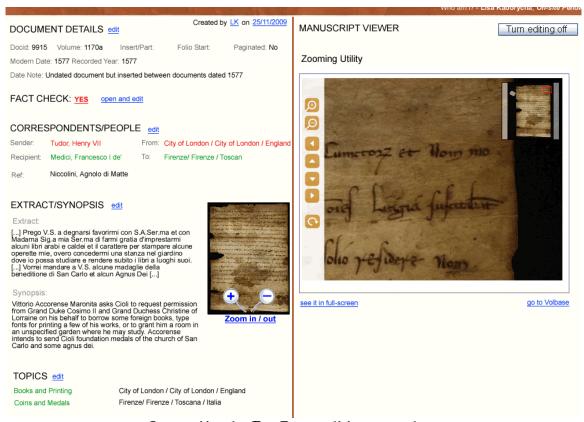


Image 11 - the DocBase, editing turned on

Each document file has the following main fields: Document Details, Correspondents/People, Extract/Synopsis, and Topics. Each main field has its sub-fields except for Topics

Document Details

Docid

Volume

Insert/Part

Folio start

Paginated

Modern Date

Recorded Year

Date notes

Correspondents/People

Sender Sender from Recipient Recipient from Reference to

The Reference to field is a list of people mentioned in document's extract [see 2.4.3].

Extract/Synopsis

Extract Synopsis

The extract is the actual transcription of the document. The synopsis is a summary of the transcription; it is written in English [see 3.4 for details].

Topics

Topic Topic Place

Here the user can add one or more of the MAP topics appropriate to the document, along with the place or places associated with it [see 2.4.4].

2.4.2 VolBase sub-module

This is the data entry part that refers to the volumes of the Medici Granducal Archival Collection.

Using this interface the scholar can edit the volume information, but she can also flip through the folios of the digitized volume [see 2.8].

By clicking on a digitized folio, she can, if already indexed by the Fellows, access its transcription. Otherwise, if she has permission to do so, the Fellow can enter that indexed folio's transcription, creating a new database record.

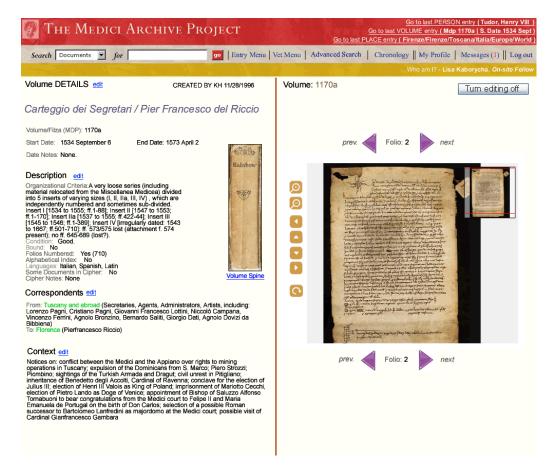


Image 12 - the VolBase, editing turned on

Each volume has the following main fields: Volume Details, Description, Correspondents and Context.

Each main field has its sub-fields except for Context:

Volume Details

Mdp or Volume Number Start Date End Date Date Note

Description

Organizational Criteria
Condition
Bound
Folios Numbered
Alphabetical Index
Languages
Some Documents in Cipher
Cipher Notes

Correspondents

From To

Context

This section provides a brief historical overview of the specific events that are featured in the volume.

Volume Spine

If the volume spine has been digitized, the image of it is shown here.

2.4.3 PeopleBase sub-module

This is the data entry part, which stores biographical information regarding the individuals and corporate bodies whose names appear in the DocBase.

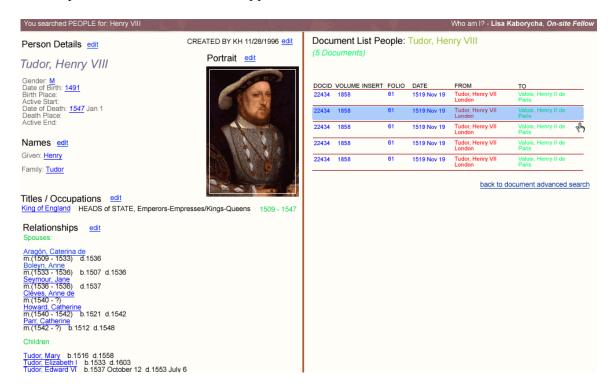


Image 13 – the PeopleBase, editing turned on

Each bio file has the following main fields: Person Details, Names, Titles and Occupations, Relationships and Portait. Each main field has its sub-fields except for Names, Titles and Occupations and Portrait.

Person Details

Gender

Date of birth

Birth place (if the place is not present in the database the user can click a button called "add new place record")

Active start

Date of Death

Active end

Names

This is a list of name variations of the same person that the Fellows encounter and that they have annotated during their work. It is used mainly for search purposes.

Titles and Occupations

This is a list of the titles and occupations that the person has obtained during his/her life with the dates if they are known.

Relationships

Spouses Children

Portrait

This is the portrait of the person. If the image box is empty, anyone can upload a portrait image. The Community Coordinator will be notified if this happens and will decide whether or not to allow the image to be updated and therefore viewed by the users. If an On-Site Fellow or a Senior Distance Fellow has uploaded the image, it automatically goes online.

2.4.4 GeoBase sub-module

This is the data entry part, which contains the hierarchy of places used in the database from continents and nations down to individual buildings, institutions and sites.

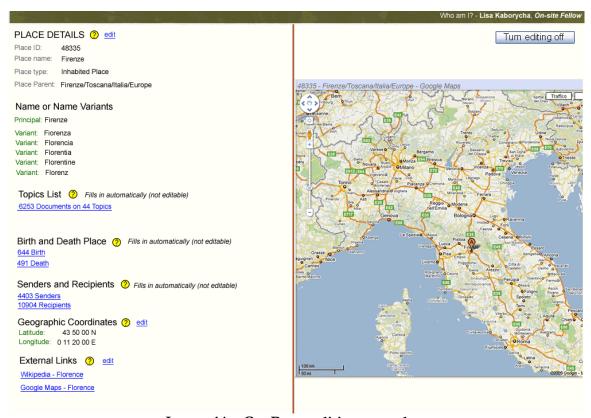


Image 14 - GeoBase, editing turned on

Each Place file has the following main fields: Place Details, Names or Name Variants, Topic List, Birth and Death Place, Senders and Recipients and External Links.

Place Details

Place ID

Place name

Place type

Place parent (the TGN format field for example Firenze/Firenze/Toscana/Italia, means placerecord/province/region/country).

Name or Name Variants

This is a list of the names the user can find in the documents when somebody is referring to this place.

Topic List

This is a not-editable field. It fills in automatically, listing the number of topics assigned to in the place file. If clicked, it shows the list of the documents that refer to that place. This is done using the split-screen feature [see 3.3].

Birth and Death Place

Birth Death

This is a not-editable field. It fills in automatically and gives both the number of *Birth place* and *Death place* referred to in the place file. When the user clicks one of these links, he gets a list of persons who were born and/or died in that place in the right screen. This is done using the split-screen feature [see 3.3].

Senders and Recipients

Senders Recipients

This is a not-editable field. It fills in automatically and gives both the number of Senders and Recipients referred to in the place file. When the user clicks one of these links, a list of documents sent from, or sent to, that place appears in the right screen. This is done using the split-screen feature [see 3.3]. Both document links are clickable as well as senders and recipients, so the user can select a document file or a bio file.

Geographic Coordinates

Latitude Longitude

These are used to create the Google Map that shows up in right screen. If the place is registered in the TGN, the geographic coordinates are automatically given. Otherwise, the Fellow has to find and enter them manually.

External Links

This is a list of links to external sites. It should include a link to show the place on Google Maps and can include a Wikipedia page as well as other useful links that the Fellows will attach.

Right Screen (default)

The default right screen displays the place record on Google Maps.

This is obtained using the Google Maps APIs [see 6.7]. The database table tblPlaces will contain a new field for the geographic coordinates of the place record from which the Google Maps APIs will create the appropriate Google Map [see 4.3.6].

2.5 Community Module

Each user of the system (excluding the Guest user group, see 2.1.1.7] is part of The Medici Archive Project Scholarly Community.

A Welcome Page welcomes a user login:



Image 15 – the Welcome Page

From the Welcome Page the user can access every part of the system using the Main Menu [see 3.2] or can click on any of the "From your last logon" links.

This page displays a link to the MAP teaching platform http://courses.medici.org in order to inform Community Users of our current curriculum offerings. We expect to recruit many interested and qualified future online scholars through this channel.

The Community module consists in three sub-modules:

- Comments
- Profiles
- Personal Messages

2.5.1 Comments

Each database item (volume, document, bio or place) can be commented on using a special section on the page called "View Comments - Add Comments" (see image below). Community users can comment on either vetted or non-vetted documents. Comments made prior to vetting, will be distinguished from the post-vetting comments and will show the "pre-vetting" label. The Community Coordinator [see 2.1.1.8] is in charge of moderating the posted comments.

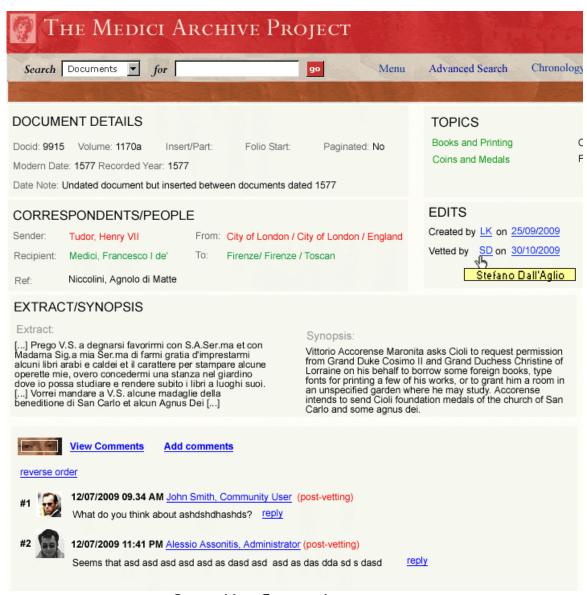


Image 16a – Community comments (post-vetting comments and Edits section)

2.5.2 Profiles

Each user has his own profile, which is his personal business card for the community:



Image 16b – Community, User's Profile



Image 16c – The user is looking at another user's profile

These are the profile fields:

First name and Last name A user's picture (the user can upload a picture of himself to the system) *Group* (the system's user group he belongs to) Title (the title the user has at the organization where he works) Organization

Location

Interests

Resume (the user can upload a pdf or word file to the system)

All of these fields are editable except for *Group*, which is assigned by the Authentication module [see 2.1]. While accessing another profile page, a user can send a personal message to the user to whom the profile belongs [see 2.5.3].



Image 16d – Profile Page with the Help system and the MAP Courses link

Aside from the user's profile itself, in this page it is possible to access MAP's online teaching platform (under the section MAP Courses), and the online manuals [see 5.1].

All the profile data is stored in *OpenLDAP* [see 6.1.2].

2.5.3 Community Messaging System

2.5.3.1 Personal Messages

Each user has his own inbox where he receives personal messages and can reply to those he is interested in:

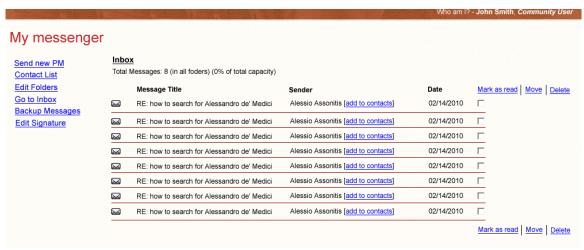


Image 17 – Community, A User's Inbox

2.5.3.2 System Messages

If the user has a special role (Distance Fellows Coordinator, Community Coordinator or Digitization Coordinator), he receives system messages in his inbox. A Distance Fellows Coordinator can receive a notification about a new document entered by a Distance Fellow; a Community Coordinator can receive a notification of a new comment posted that is waiting for approval; a Digitization Coordinator can receive the notification that a new digitized volume is ready. Also *On-Site Fellows* receive notifications about documents that have been finished and set to "ready to be vetted" [see 3.2.1] by *Distance Fellows*. Users with special roles can also send messages to an entire user group.

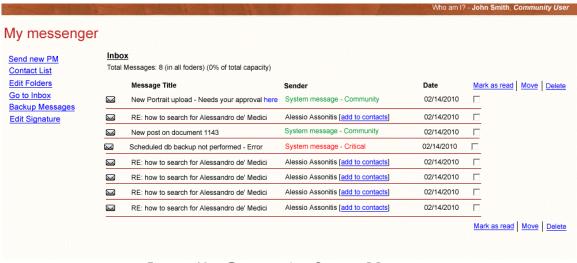


Image 18 – Community, System Messages

2.6 Administration Module

The Administration is used both to administer the whole system (user account management, reports, statistics, logs, document revisions) and to manage the data backups.

2.6.1 User Management

Through the *User Management* dropdown menu, users belonging to the *Administrators* user group will be able to activate new user accounts, create new user accounts, and modify existing ones.

This is the dropdown structure of the menu:

Browse users Activate user accounts Add new user Edit user

This menu is shown at section 3.2.3.

2.6.2 Reports and Revisions

This menu helps the administrative user to browse the data that the History Log Module recorded [see 2.2].

Through the Reports and Revisions dropdown menu the Administrator will be able to search through the logs [see 2.2.1] and the revisions [see 2.2.2] using several types of filters.

This is the dropdown structure of the menu:

Activity Report
Comments Report
Statistics
Revisions Database
Document revisions
People revisions
Volumes revisions
Places revisions

The Activity Report helps the Administrator to browse the user's logged activities.

The *Comments Report* helps the Administrator to track the comments posted by the users and all the activities that regard the community itself.

The *Statistics* tool is able to perform statistical graphs using each type of data recorded by the history log system combining it with the database content.

The Revisions Database is a search tool that will be able not only to search for previous revisions of documents, volumes, bios and/or locations but it can also restore an old version of it.

This menu is shown at section 3.2.3.

2.6.3 System Management

Through the *System Management* menu the user will be able to see the server status and to access the backup system.

This is the dropdown structure of the menu:

Servers' status
Backup and Restore
Scheduled Backup
Manual Backup
Browse Data Backups
Browse Logs Backups

The *Servers' Status* takes the Administrator to a page that shows all the sensitive server information such as: cluster information, uptime of all the servers connected, disk use, load average etc.

The Backup and Restore tool is used to manage the backup system.

The *Scheduled Backup* is a tool to be used to schedule automatic backups. It permits you to define what to backup and when to do it.

The Manual Backup is a tool to perform a fast backup in a few clicks by selecting the data you want to backup.

The *Browse Data Backups* and the *Browse Logs Backups* are two different pages to search through data backups or logs backups.

This menu is shown at section 3.2.3.

2.7 Digitization module

The digitization of the manuscripts is the most important task to be performed along with the development of the software, so that when MAP launches the production release, the Distance Fellows and the Community users will have a large number of digitized manuscripts to work on and discuss.

The digitization tasks are performed separately from the On-site Fellows' work. This means that the Fellows' work is independent from the digitization task; the On-Site Fellows in fact have the primary sources to look at. Conversely, the process of digitization is not dependent on the creation of database records (transcriptions).

2.7.1 The digitization task

The Medici Granducal Archival Corpus is organized in volumes (called *filze*). At first the digitization technician "prepares" the volume: she analyzes the volume's conditions, deciding if it has to be restored, she compiles a catalog description [see *Minutes of Meeting 3 – Catalog Description*] and eventually numbers the pages.

At this point she starts taking pictures of each page including the spine; at the end of this process she creates a directory with the name of the volume (ex. 1701a) on his computer and she transfers all the images produced from the digital camera to the computer.

- progressive-number: 4 numbers that identify the page univocally (ex. 0001,0023).
- *type*: can take one of the following values: *COPERTA*, *C* (carta), *G* (guardia) e *A* (allegato).
- page-number: 3 numbers that identify the page number (ex. 001,009).
- side: can take one of the following values: R (recto), V (verso).

Examples:

0001_COPERTA_001_R.tiff 0456_C_001_R.tiff 0456_C_001_V.tiff

The directory is afterwards compressed in a zip format and sent to the MAP system using the digitization interface [see 2.7.2]¹.

¹ The digitization task will start at the same time as the development process so MAP will not be able to provide the digitization interface. The digitized manuscripts will be stored on dvds and/or hard drives and when the digitization interface is ready they will be uploaded to the system. These images will also help us for the initial testing of the digitization module.

For easy access of a document in its digital format, the digitized manuscripts are stored in the database in a database table called *tblDigitalDocs* [see 4.3.5].

A new record is created for each new digitized document uploaded. Each database record that belongs to the table *tblDigitalDocs* corresponds to two digitized images, *recto* and *verso*.

The digitized documents database table does not contain any external keys to the *tblVolumes* or to the *tblDocuments* tables [see 4.2]. This database table, however, includes all the information, extracted from the names of the files and from the directory in which it resides, allowing the association to a volume and to the folio (page). The On-Site or the Distance Fellow will manually link the extract to the correct folio or folios to which it belongs.

The system performs these actions:

- 1. Checks if the volume is already present in the system
- 2. If it is, asks if the user wants to replace the old content with the new one
- 3. Saves the original zip file in the digitization root folder and renames it in this format: date_volumenumber.zip
- 4. Decompresses the zip file in a folder (renaming it with the volume number)
- 5. Analyzes the names of the files
- 6. Creates a Tiled Pyramidal TIFF [see 2.8]
- 7. Creates a new record in the digitization table, *tblDigitalDocs* [see 4.3.5]

2.7.2 The Digitization web interface

The Digitization module consists in the web interface the Digitization Technician uses for uploading the digitized volumes to the system and to compiling the catalog descriptions.

For a detailed explanation of the *Digitization web interface* see section 3.2.4.

2.8 Manuscript Viewer

To show the digitized manuscripts, MAP chose to adapt and to integrate IIPImage in the web application. IIPImage is an open source lightweight streaming client-server system for the web-based viewing and zooming of ultra high-resolution images. It is designed to be bandwidth and memory efficient and usable even over a slow Internet connection. To use IIPImage though, the source images must be in a multi-resolution tiled pyramidal format. Tiled Pyramidal TIFF is a tiled multi-page TIFF image, with each resolution stored as a separate layer within the TIFF. During the digitization process the system creates the Tiled Pyramidal TIFF using Imagemagick².

2.8.1 MAP's IIPImage implementation

The manuscript viewer will be available in two different formats: *full-screen* and *half-size*, and will be accessed from both *VolBase* [see 2.4.2] and *DocBase* [see 2.4.1].

2.8.1.1 Full-screen

This format is normally used by a Fellow when working in the *VolBase* [see 2.4.2]; she flips through a volume of the Medici Granducal Archival Collection and reads each page looking for a document that matches one of MAP's forty-four topics.

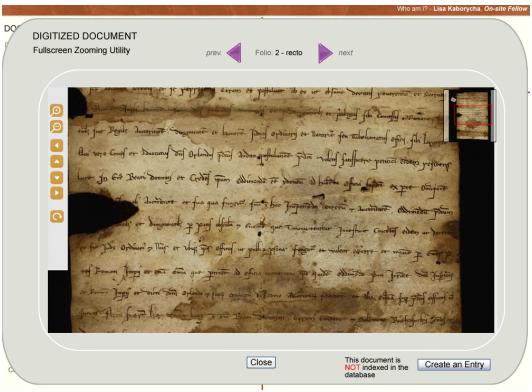


Image 19 - Manuscript Viewer (Full-screen)

-

² ImageMagick is an open source and free software suite to convert and write images in a variety of formats, including TIFF.

The two violet arrows at the top of the document allow the user to flip through the document pages.

The button at bottom right allows the user to create a new document record, if the document is not already present in the database. If it is, the button will have the "Show me this Document" label instead of "Create an Entry".

2.8.1.2 Half-size

The half-size format is used to show the user a portion of the manuscript while reading transcriptions and of the document or the volume details [see *Image 10* and *Image 11* in section 2.4.1 and 2.4.2].

It is also used to transcribe the document:

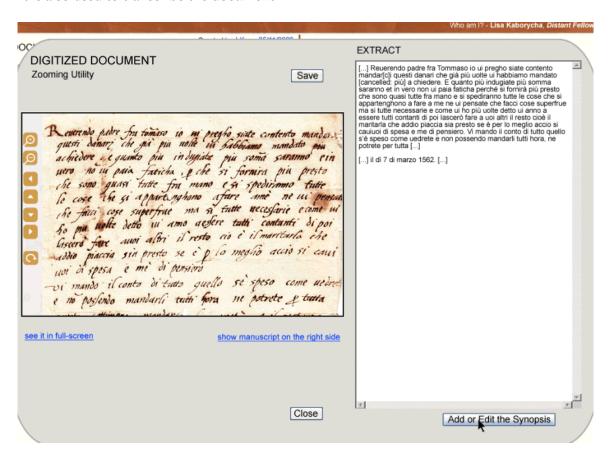


Image 20 – Manuscript Viewer (Half-size)

2.9 Deployment Diagram

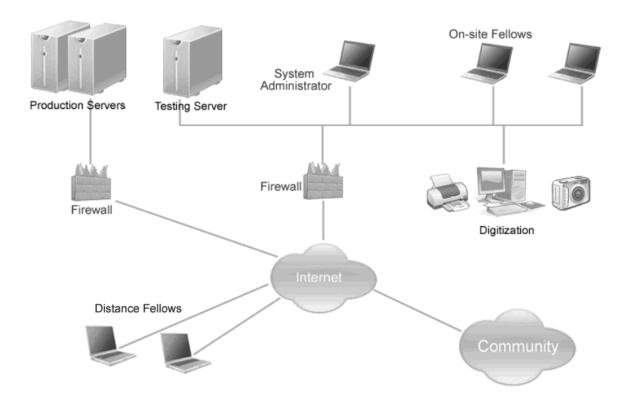


Image 21 - Deployment Diagram

Both On-Site and Distance Fellows use a web browser to enter the data into the system. During their work they perform frequent database queries and they can access the digitized documents and volumes.

The Digitization Technician digitizes documents on his workstation and then uploads them to the server using the *Digitalization web interface* [see 3.2.4].

The System Administrator does basic maintenance operations and checks the logs. Updates and bug-fixings are done by logging directly onto the server through SSH.

Community Users access the system using the web browser. They view all content, and all digitized images. They can comment on each database item (document, bio, volume or place) and read the other users' comments. They can also send personal messages to the other community members and receive answers.

The entire system resides on two servers. For servers' deployment details see section 8.1.

3. User Interface Design

3.1 Usability

MAP would like to preserve the usability of the current interface, associating to each of the four data entry sections to four different colors:

- the DocBase [see 2.4.1] is identified by a brown navigation bar



Image 22 (a,b,c,d) - NavBars

3.2 The Main Menu

The Main Menu is used to navigate through the interface. Each user group has its own menu.

3.2.1 On-Site Fellows and Distance Fellows menu

It shows six links to perform different actions, a search form (used for the Basic Search see 2.3.1] and the *Shortcut menu* [see 3.2.1.9] at top right.



Image 23 - Main Menu (Onsite Fellow or Distance Fellow)

This module helps the user to perform several actions:

3.2.1.1 The Search form (Basic Search)

It is a form for the user to perform a simple word search on documents, volumes, people, locations and community comments [see 2.3.1].



Image 24 – Basic Search

3.2.1.2 Entry Menu

The entry menu is visible to Administrators, On-Site Fellows, Senior Distance Fellows and Distance Fellows. It is used to perform the "Add" operations and to delete the item the user is currently viewing.



Image 25 - Entry Menu

Menu items:

Add new document

Add new volume

Add new person

Add new place

Delete this item (visible to Administrators and On-Site Fellows)

3.2.1.3 Vet Menu

The Vet Menu (vetting menu) is visible if we are editing a document in the DocBase. A Fellow can set it as "Ready to be vetted" to permit another Fellow to do a check and review it. After the second Fellow (who must be an On-Site Fellow or a Senior Distant Fellow) has checked and corrected it, she will vet the document using the "Vet this document" option. This menu is also used to un-vet the document in case someone notices an error. In this case the document has to be re-vetted by another Fellow.

If a document is not set as "Ready to be vetted" it is allowed to be viewed and edited only by the user who created that entry.



Image 26 – Vet Menu

Menu items:

Set Ready to be vetted

Vet this document (visible to Administrators, On-Site Fellows and Senior Distant Fellows if the document is "ready to be vetted")

Un-vet this document (visible to Administrators, On-Site Fellows and Senior Distant Fellows if the document is "vetted")

3.2.1.4 Advanced Search

The Advanced Search menu is used to do a search over documents, volumes, people and places, and community comments applying several filters [see 2.3.2].



Image 27 - Advanced Search Menu

Menu items:

Search on documents

Search on volumes

Search on persons

Search on places

Search on community comments

3.2.1.5 Chronology

The Chronology dropdown menu will show a list of the twenty most recent database items viewed or edited.

3.2.1.6 My profile

This link takes the user to his profile [see 2.5.2].

3.2.1.7 Messages

This link takes the user to his inbox [see 2.5.3].

3.2.1.8 Log out

This link logs out the user from the system.

3.2.1.9 Shortcut Menu

The Shortcut Menu is placed at the top right of the *Main Menu* and it consists in three links (taken from the chronology) that if clicked take you to the last document, person, volume or place that you most recently edited (see image below).

3.2.2 Community User menus

When a Community user logs on she can perform the same actions that Fellows can [see the section above), with the exception that she does not have access to the *Entry menu* and the *Vet Menu*, and therefore cannot edit or delete anything.

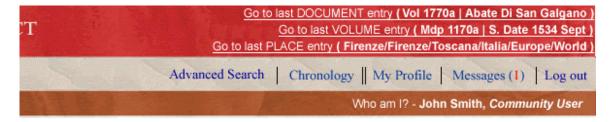


Image 28 - Community User - Main Menu

3.2.3 Administration menus

When a user belonging to the Administrators user group logs on, the Administration menus appear. This module consists in three dropdown menus: *User Management, Reports and Revisions*, and *System Management*:

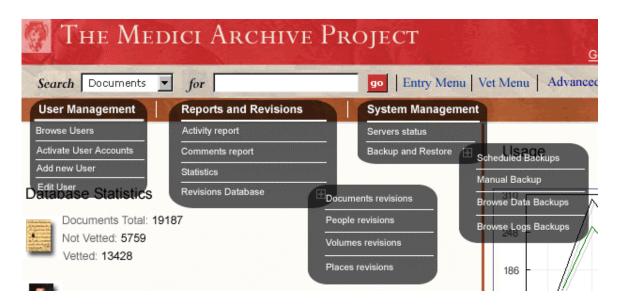


Image 29 - Administration menu

3.2.4 Digitization menu and its web interface

Once the Digitization Technician [see 3.1.1.5] logs on, she is forwarded to the Digitization Interface that consists in two main menus: *Digitized volumes/filze upload* and *Catalog Descriptions*.

This menu helps the Digitization Technician to perform two main actions:

- Prepare the volume compiling its catalog description [see *Minutes of Meeting 3 Catalog Description*], which can be downloaded as a pdf file to be handed to the State Archive official.
- Upload the entire digitized volume: all the images are prepared before the upload in the correct format following the State Archive cataloguing conventions [see *Minutes of Meeting 3 State Archive in Florence requirements*] and compressed into a zip file.

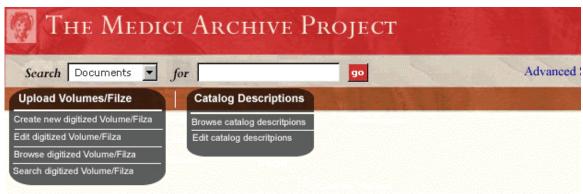


Image 30 – Digitization menu

This is the dropdown structure of the menu:

Create new digitized volume/filza Edit digitized volume/filza Browse digitized volumes/filze Search for volume/filza

Create new volume/filza allows the user to compile a new catalog description of the volume and then to upload the digitized images to the system.

Edit volume/filza allows the user to edit a previous catalog description and/or allows uploading of a different set of images for that volume.

Browse volumes/filze allows the user to look at a list of digitized volumes.

Search for volume/filza allows the user to perform a search through digitized volumes.

3.2.4.1 Catalog descriptions

This menu helps the Digitization Technician to find and download the catalog description to be handed to the State Archive officials.

This is the dropdown structure of the menu:

Browse catalog descriptions Edit catalog descriptions

Browse catalog descriptions allows the user to get a list of the catalog description of each volume and lets them download a pdf version of each catalog description.

Edit catalog descriptions allows the user to search and edit a catalog description without needing to use the Edit digitized volumes/filza page.

To see the fields listed in any catalog description, see the document *Minutes of Meeting 3 – Catalog Description*.

3.3 Split Screen

This feature is one of the most important innovations of the whole system. The Fellows, in fact, requested a way to compare one document to another entry in the same screen. To do this MAP will create an interface that is divided into two parts. The left-hand side of the screen will be where the Fellow modifies the document. On the right-hand side the Fellow can perform a related search. Here, the Fellow will also be able to open up a "read-only" version of the document, people, volume or place entry for comparison, while editing on the left-hand side of the screen:

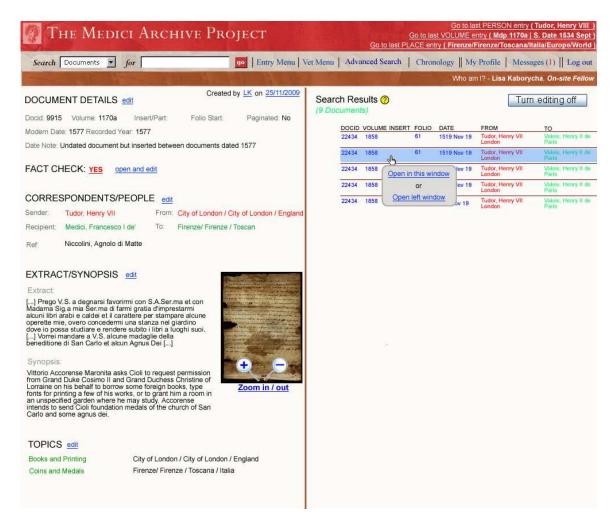


Image 31 – Left screen (Edit Screen) and Right screen (Searches – Read-only Screen)

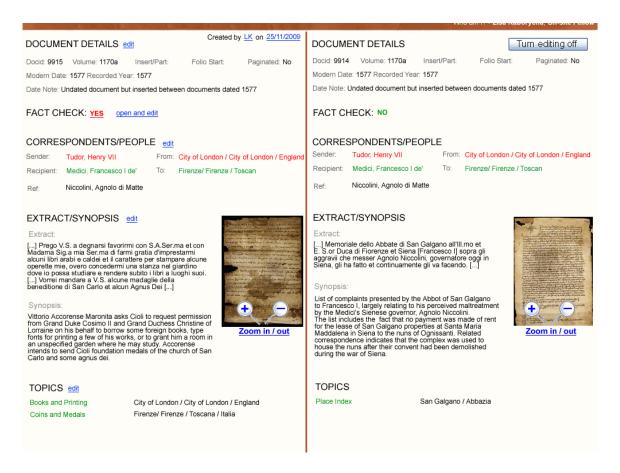


Image 32 – Left Side (Edit Screen) and Right Side (Document Read-only Screen)

Above you can see a Fellow that is comparing the document she is editing to another readonly document. Here on the other hand, a Fellow is comparing a document to a person record, which is mentioned in the document.

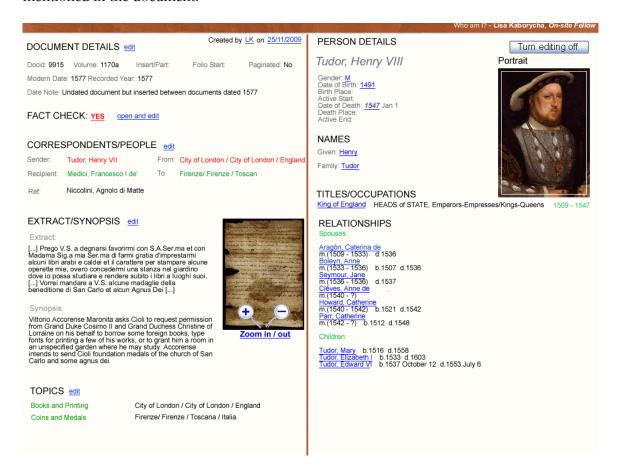


Image 33 – Left Side (Edit Screen) and Right Side (Person Read-only Screen)

3.4 The Document Transcription

The document transcription task is one of the core actions that the Fellows perform, and one of the main goals MAP wants to accomplish is performing the document transcriptions with the use of digitized documents. Therefore, it is very important to provide the Distance Fellows with usable and accessible workbench.

Clicking the Extract/Synopsis section edit link the Fellow enters the transcription tasks:

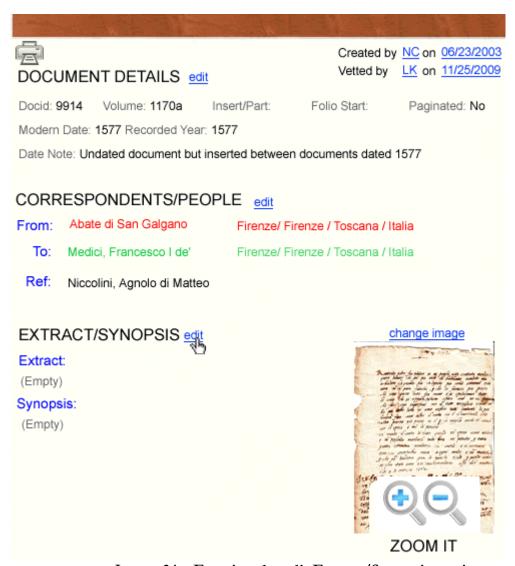


Image 34 – Entering the edit Extract/Synopsis section

The "Transcription window" is divided into two parts: on your left side the manuscript viewer is shown, and on the right side the Fellow can enter data in the Extract. The Fellow is able to put the Manuscript viewer on the right side and the extract page on the other side by clicking the "show manuscript on right side" link.

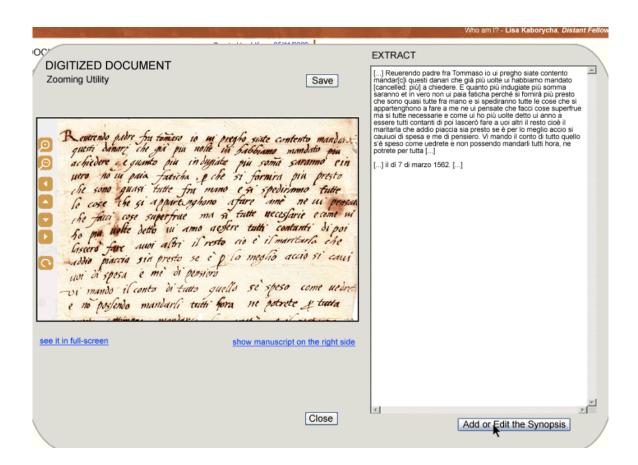


Image 35 – The Transcription window

After transcribing the document, the Fellow clicks on the button "Add or Edit Synopsis".

The previous screen slides on the left and the user starts entering the synopsis comparing it to the previously entered extract:

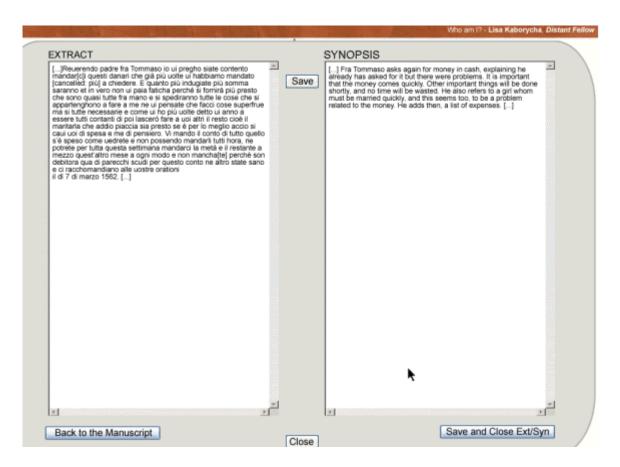


Image 36 - The Synopsis Window

When the Fellow has finished she presses the button labeled "Save and Close Ext/Syn".

The Extract/Synopsis section of the DocBase has been filled.



Docid: 9914 Volume: 1170a Insert/Part: Folio Start: Paginated: No

Modern Date: 1577 Recorded Year: 1577

Date Note: Undated document but inserted between documents dated 1577

CORRESPONDENTS/PEOPLE edit

From: Abate di San Galgano Firenze / Toscana / Italia

To: Medici, Francesco I de' Firenze / Toscana / Italia

Ref: Niccolini, Agnolo di Matteo

EXTRACT/SYNOPSIS edit

Extract:

[...]Reuerendo padre fra Tommaso io ui pregho siate contento mandar[c]i questi danari che già più uolte ui habbiamo mandato [cancelled: più] a chiedere. E quanto più indugiate più somma saranno et in vero non ui paía faticha perché si fornirà più presto che sono quasi tutte fra mano e si spediranno tutte le cose che si appartenghono a fare a me ne ui pensate che facci cose superfrue ma si tutte necessarie e come ui ho più uolte detto ui anno a essere tutti contanti di poi lascerò fare a uoi altri il resto cioè il maritarla che addio piaccia sia presto se è per lo meglio accio si caui uoi di spesa e me di pensiero. Vi mando il conto di tutto quello s'è speso come uedrete e non possendo mandarli tutti hora, ne potrete per tutta questa settimana mandarci la metà e il restante a mezzo quest'altro mese a ogni modo e non mancha[te] perché son debitora qua di parecchi scudi per questo conto ne altro state sano e ci racchomandiano alle uostre orationi il dì 7 di marzo 1562. [...]

Amen's fair for the same and grade and content and a state of the same and a s

change image



ZOOM IT

Synopsis:

[...] Fra Tommaso asks again for money in cash, explaining he already has asked for it but there were problems. It is important that the money comes quickly. Other important things will be done shortly, and no time will be wasted. He also refers to a girl whom must be married quickly, and this seems too, to be a problem

Image 37 – Completed Extract/Synopsis section

3.5 Getting info from dropdown auto completing menus

Another important feature, which was requested by the Fellows, is getting items' information directly from the dropdown auto completing text fields.

In the image below the user is selecting a Sender by typing in an auto complete text field: a dropdown suggestion box appears with five fields; the last field is the corresponding link to the person description.

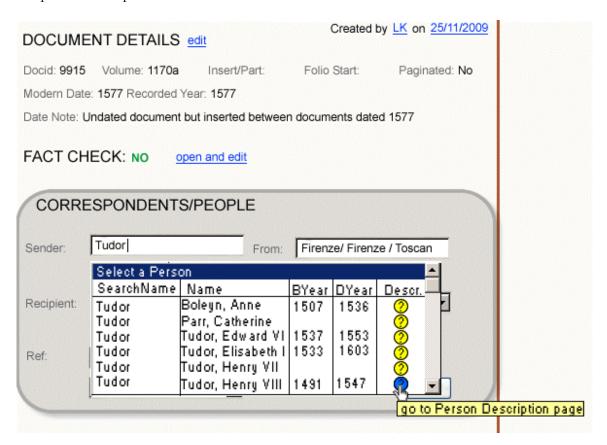


Image 38 - People suggestion box with link to PeopleBase

By clicking the "?" link, a description page of the bio selected appears. This is obtained by the use of the spit-screen feature [see 3.3]:

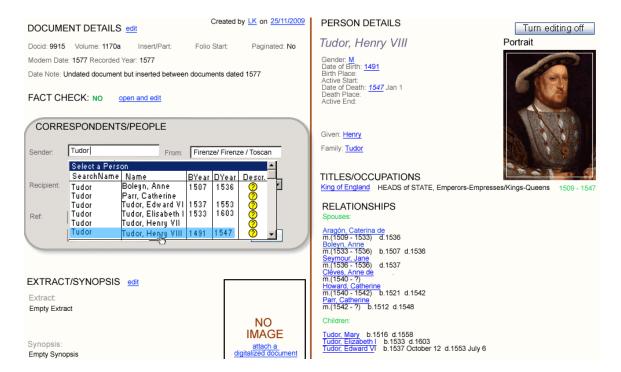


Image 39 - People suggestion box with split-screened PeopleBase

At this point the user is able to see the information related to that person and see if it is the one she was looking for.

This feature works for people, volumes and places dropdown suggestion boxes.

The same feature applies to the topic selection:

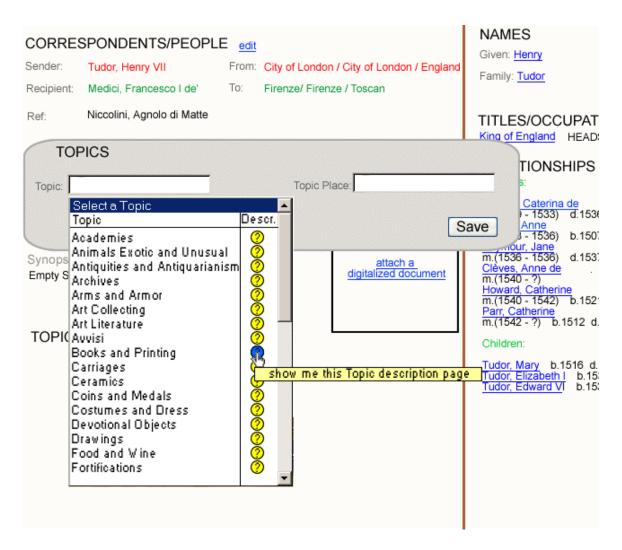


Image 40 - Topics suggestion box with link to Topic info box

In this case though, the split-screen feature is not used. The information is displayed using a floating help box. From this box you are able to directly chose the topic if it meets the users' needs, by clicking the button "Choose":

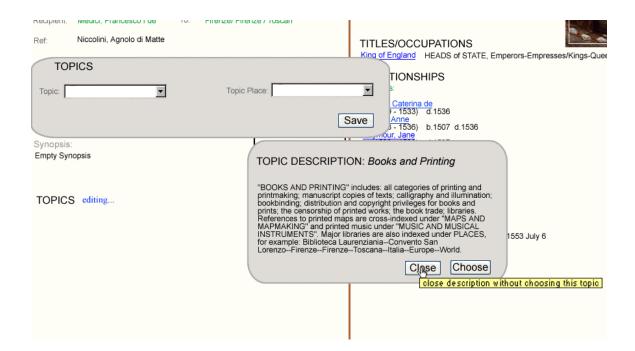


Image 41 – Topics suggestion box with link to floating Topic info box

All these features are developed using DWR and JQuery. See 6.1.1 and 6.1.5 for details.

3.6 How to print records

Another important feature, which was requested by the Fellows, is an easy-to-use printing facility for the documents indexed in the database.

3.6.1 Printing a single item page

To print a selected record the user has to click the print link at top left of the screen:

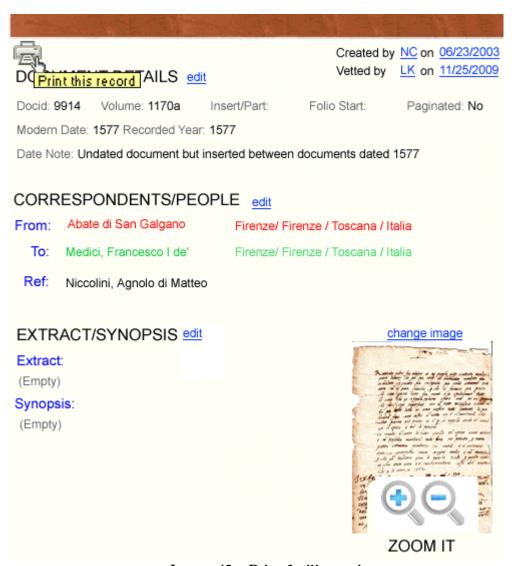
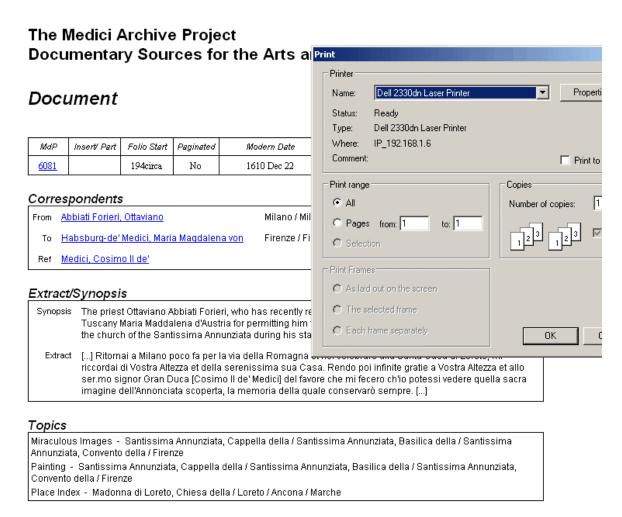


Image 42 – Print facility on item page

After clicking the link a preview page is shown:



Research Notes

No Research Notes Available.

Document Citation Text:

Archivio di Stato di Firenze, Mediceo del Principato 6081 folio 194circa.

(Entry 15693 in the the Medici Archive Project Documentary Sources database.)

The Medici Archive Project

All information cital from the Documentary Sources for the Arts and Humanities database must be credited to:

Image 43 – Print preview page

The user is then able to print it by clicking the operating system print button (in this case the Microsoft Windows).

3.6.2 Printing document records from a search list

It is also possible to select and print documents from a search list:

	cuments)	FOLIO	DATE	FROM		Selec
22434	1858	FOLIO 61	1519 Nov 19	Tudor, Henry VII London	Valois, Henry II de Paris	
22434	1858	61	1519 Nov 19	Tudor, Henry VII London	Valois, Henry II de Paris	V
22434	1858	61	1519 Nov 19	Tudor, Henry VII London	Valois, Henry II de Paris	
22434	1858	61	1519 Nov 19	Tudor, Henry VII London	Valois, Henry II de Paris	✓
22434	1858	61	1519 Nov 19	Tudor, Henry VII London	Valois, Henry II de Paris	V
22434	1858	61	1519 Nov 19	Tudor, Henry VII London	Valois, Henry II de Paris	✓
22434	1858	61	1519 Nov 19	Tudor, Henry VII London	Valois, Henry II de Paris	
22434	1858	61	1519 Nov 19	Tudor, Henry VII London	Valois, Henry II de Paris	
22434	1858	61	1519 Nov 19	Tudor, Henry VII London	Valois, Henry II de Paris	

Image 44 - Selected four documents to be printed from a search results list

First, the user selects the documents that he wants to print and then clicks the link "Print Selected". The user is also able to select all the documents by clicking the "Select All" link, and then to print them out.

3.7 Compatibility

The software produced will be web-based and therefore it will be accessible from any of the major web-browsers.

To guarantee cross-browser compatibility, the Graphic Designer along with the Developers will produce HTML according to the W3C standards.

4. Data Structural Design

4.1 Data Structure

The current database structure is preserved:

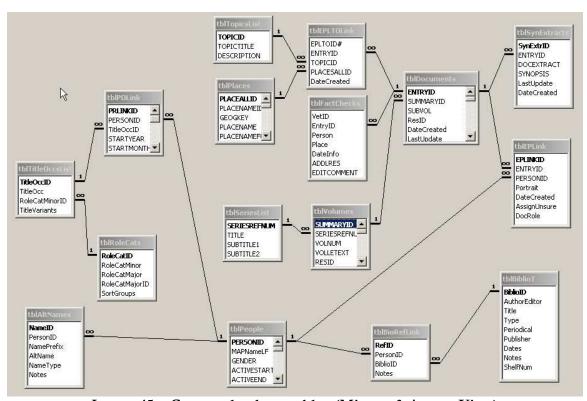


Image 45 – Current database tables (Microsoft Access View)

MAP's online database http://documents.medici.org is based on Mysql, and contains the same database tables structure of the Microsoft Access-based On-Site software.

To populate the online database, MAP at the present time exports the data from Microsoft Access to Mysql using some Perl-based scripts once a week.

These scripts will need to be revised for use with the new system, but 90% of the Mysql structure MAP now has will be used "as is".

4.2 Current Database Tables

TblSeriesList				
SeriesRefNum	Number	(primary key) Unique identifying number for Series		
		Reference titles and subtitles		
Title	Text	Series reference title, from Archivio Mediceo del		
		Principato, Inventario Sommario (Roma, 1951)		
SubTitle1	Text	Series reference subtitle 1, from Archivio Mediceo		
		del Principato, Inventario Sommario (Roma, 1951)		
SubTitle2	Text	Series reference subtitle 2, from Archivio Mediceo		
		del Principato, Inventario Sommario (Roma, 1951)		

	TblVolumes (VOLUMES)				
SummaryId	AutoNumber	Volume Summaries			
SeriesRefNum	Number	(extern key to tblSerieList) Unique identifying number keyed to SERIESLIST table containing preassigned titles of the categories into which the Mediceo del Principato archive has been divided.			
VolNum	Number	Volume Shelf Number (e.g., 1176). Will not exceed 4 digits			
VolLeText	Text	Letter extension for volume shelf number (e.g., "a" in 1176a)			
ResId	Text	Initials of Researcher who created the Volume Record			
DateCreated	Date-time	Date volume record is created			
VolTobeVettedDate	Date-time	Date Volume is signed off for vetting (whether from backlog or newly entered volume)			
VolTobeVetted	Yes/Not	Check box to signal Volume is ready for vetting			
VolVetId	Text	Initials of Researcher who vetted the volume			
VolVetBegins	Date-time	Date Vetting Begins			
VolVetted	Yes/Not	Check box to signal volume vetting is complete			
VolVettedDate	Date-time	Date volume vetting is completed			
STATBOX	Text	Vetting Status of Volume			
STARTYEAR	Number	For inclusive date range of all documents in the volume.			
STARTMONTH	Text	For inclusive date range of all documents in the volume. DELETE - USE NUMBER FORMAT WITH MONTHNAME LOOKUP TABLE			
STARTMONTHNUM	Number	Start Month Number Format			
STARTDAY	Number	For inclusive date range of all documents in the volume.			
ENDYEAR	Number	For inclusive date range of all documents in the			

		volume.
ENDMONTH	Text	For inclusive date range of all documents in the volume. DELETE - USE NUMBER FORMAT WITH MONTHNAME LOOKUP TABLE
ENDMONTHNUM	Number	Start Month Number Format
ENDDAY	Number	For inclusive date range of all documents in the volume.
DATENOTES	Memo	For additional information characterizing date range of documents. E.g., "Document dating is a mix of Florentine and Roman reckoning."
SENDERS	Memo	General description of correspondents. E.g., From:Tuscan Court (Averardo Serristori; Cosimo I de' Medici; Francesco I; Cardinal Ferdinando [I] de' Medici)
RECIPS	Memo	General description of correspondents. E.g., To: Rome (Ambassadors, Agents, Administrators) and the Papal Court (Pope Paul III).
CONTEXT	Memo	General historical description of the volume emphasizing major events. E.g., ""Pope Paul III elevates new cardinales; Medici and Ferrarese ambassadors' precedence controvery at the Papal Court; Reports on plague in Rome."
FOLIOCOUNT	Text	Number of folios in the volume, including blank folios. Based on pagination. Folio count in unpaginated volumes are estimated.
BOUND	Yes/No	Check box to signal that volume is bound or unbound.
FOLSNUMBRD	Yes/No	Check box to signal that folios are paginated or unpaginated.
OLDALPHAINDEX	Yes/No	Check box to signal presence of an index. Some volumes contain old alphabetical indices of names.
CONDITION	Memo	Notes on physical condition of the volume and individual folios. E.g., ""Documents heavily treated and sometimes faint. Deterioration sometimes affects legibility."
ITALIAN	Yes/No	Check box to signal presence of documents in ITALIAN language.
SPANISH	Yes/No	Check box to signal presence of documents in SPANISH language.
ENGLISH	Yes/No	Check box to signal presence of documents in ENGLISH language.
LATIN	Yes/No	Check box to signal presence of documents in LATIN language.
GERMAN	Yes/No	Check box to signal presence of documents in GERMAN language.
FRENCH	Yes/No	Check box to signal presence of documents in FRENCH language.

OTHERLANG	Text	Text field to list presence of documents in
		additional languages not included in the check box
		option
CIPHER	Yes/No	To signal presence of documents in cipher
CIPHERNOTES	Text	Text field for descriptive notes on documents in
		cipher. E.g., "Names and brief passages in code,
		sometimes transcribed."
ORGNOTES	Memo	Volume organization. E.g., ""5 inserts of unbound
		documents; pencil numbering in first four inserts;
		insert 5 unnumbered, contains various unsigned
		notes regarding Giovanna d'Austria."
STAFFMEMO	Memo	Notes for additional research required for the
		volume summary. E.g., "Identify son of Isabella de'
		Medici and Paolo Giordano Orsini, born in 1572."

	tblDoo	cuments
EntryId	AutoNumber	(primary key) Unique identifying number for Volume Entries
SummaryId	Number	(extern key to tblVolumes) Unique identifying number for Volume Summaries
SubVol	Text	Letter extension for volume shelf number (e.g., 1176a) DELETE THIS FIELD IF MATCHES WITH VOLSUM FIELD
ResId	Text	Researcher initials for entry creator
DateCreated	Date-time	Creation date of the record
LastUpdate	Date-time	Date Last Updated or Changed
DocTobeVetted	Yes/Not	Check box to signal document is ready for vetting
DOCTOBEVETTEDDATE	Date-time	Date Doc is signed off for vetting
DOCVETID	Text	ID of the Researcher who vetted the volume
DOCVETBEGINS	Date-time	Date Vetting Begins
DOCVETTED	Yes/Not	Checkbox signals record has been vetted
DOCVETTEDDATE	Date-time	Date Vetting is Completed
DOCSTATBOX	Text	Vetting Status of Document
NEWENTRY	Yes/Not	Check box to signal New Entry requiring vetting
INSERTNUM	Text	Folders containing unbound documents, usually contained within a larger box or tied together.
INSERTLET	Text	Insert Letter of folders containing unbound documents, usually contained within a larger box or tied together. "Yes" is entered for unnumbered or unlettered folders.
FOLIONUM	Number	Folio number of cited document. Left blank for unpaginated folios.

FOLIOMOD	Text	Text field For ""bis"", ""ter"", etc., modifying
0.03 700 70.0	7.7 /3.7	folio number
CONTDISC	Yes/Not	Check Box to signal discussion continued on
ID ID A CED	37 /3 T	later folio in this volume.
UNPAGED	Yes/Not	Check Box to signal unpaginated folio.
DOCDAY	Number	Day Date for Document
DOCMONTHNUM	Number	Month Number
DOCYEAR	Number	Year Date for Document
SORTABLEDATE	Text	Concatenated Year/Mon/Day for sorting purposes
YEARMODERN	Number	Year Date according to modern calendar, applied to documents dated by earlier reckoning systems. E.g., The Florentine New Year began on 25 March. Hence, a Florentine document dated 24 March 1544 is dated according to modern calendar to 24 March 1545.
RECKONING	Yes/Not	Check Box to signal that the researcher considers the calendar in use is uncertain.
UNDATED	Yes/Not	Check Box to signal undated documents.
DATEUNS	Yes/No	Check Box to signal that the document date is uncertain for any reason including illegibility, physical loss, etc.
DATEAPPROX	Text	Text field to insert researcher's estimate of approximate date assessed from document context.
DATENOTES	Text	Text field for notes concerning document date. E.g.,"1550 noted on back of letter in second hand, probably Riccio's."
SENDID	Number	Sender's Name Link containing PERSONID number to link to PEOPLELIST table.
SENDUNS	Yes/No	Check Box to signal Sender's Name is uncertain for any reason including illegibility, physical loss, name identification.
SENDLOCPLALL	Number	Sender Location ID# linked to unique ID in GETTYPLACES
SENDLOCUNS	Yes/No	Check Box to signal Sender's Place is uncertain for any reason including illegibility, physical loss, place identification.
SENDNOTES	Text	Text field for notes modifying or qualifying Sender's Name or Location. E.g.,"Giorgino Pittore Roma" or "A Niccolini Ambasciatore a Roma"
RECIPID	Number	Recipient's Name Link containing PERSONID number to link to PEOPLELIST table.
RECIPUNS	Yes/No	Check Box to signal Recipient's Name is

		uncertain for any reason including illegibility,
		physical loss, name identification.
RECIPLOCPLALL	Number	Recipient Location ID# linked to unique ID
		in GETTYPLACES
RECIPLOCUNS	Yes/No	Check Box to signal Recipient's Place is
		uncertain for any reason including illegibility,
		physical loss, place identification.
RECIPNOTES	Text	Text field for notes modifying or qualifying
		Recipient's Name or Location. E.g.,
		"Giorgino Pittore Roma" or "A Niccolini
		Ambasciatore a Roma"
GRAPHIC	Yes/No	Check Box for visual material (e.g. drawings,
		plans, coats-of-arms) etc.

tblPeople				
PERSONID	AutoNumber	Unique Identifier for a Person		
MAPNameLF	Text	Created automatically as Name Parts are Edited -		
		Full Name, Last Name First		
GENDER	Text	Gender M, F (or X for unknown or		
		placeholders)		
ACTIVESTART	Text	If no birthdate, date known activity starts		
ACTIVEEND	Text	If no deathdate, date known activity ends		
BMONTHNUM	Number	Birth Month in Number Format		
BDAY	Number	Birth Day		
BYEAR	Number	Birth Year		
BPLACEID	Number	BirthPlaceID - links to tblPlaces record		
BPLACE	Text	Birth Place - entered here only if not in tblPlaces		
		- when identified and present in tblPlaces this		
		entry will be deleted		
DMONTHNUM	Number	Death Month in Number Format		
DDAY	Number	Death Day		
DYEAR	Number	Death Year		
DPLACEID	Number	DeathPlaceID - links to tblPlaces record		
DPLACE	Text	Death Place - entered here only if not in		
		tblPlaces - when identified and present in		
		tblPlaces this entry will be deleted		
FIRST	Text	First or Given Name		
SUCNUM	Text	Succession number		
Middle	Text	Middle/Pre-Last Name		
midprefix	Text	Prefix to Middle/Pre-Last Name		
Last	Text	Last or ID Name		
lastprefix	Text	Prefix for Last or ID Name		
Postlast	Text	Name following ID Name		
postlastprefix	Text	Prefix for Name following ID Name		

BAPPROX	Yes/No	Birth Year Approx? Y/N
BDATEBC	Yes/No	Birth Year B.C.? Y/N
BPLACEUNS	Yes/No	Birth Place Unsure? Y/N
DAPPROX	Yes/No	Death Date Approx? Y/N
DYEARBC	Yes/No	Death Year B.C.? Y/N
DPLACEUNS	Yes/No	Death Place Unsure? Y/N
STATUS	Text	Active or Inactive
BIONOTES	Memo	General biographical Notes
STAFFNOTES	Memo	Internal Staff Notes
PORTRAIT	Yes/No	Portrait Description for this person Y/N - calculated from EPLINK Portrait field
FATHERID	Number	Father's Person ID
MOTHERID	Number	Mother's Person ID
RESID	Text	Researcher ID for record creator
DateCreated	Date-time	Date Entry Started on this Record
LastUpdate	Date-time	Date of Last Change

tblPlaces				
PLACEALLID	AutoNmber	Unique MAP ID (Primary key)		
PLACENAMEID	Number	Unique TGN Place (Term_ID) ID.		
		Necessarily missing for MAPPLACES and		
		MAPSITES. Must set Duplicates OK because		
		of the MAP blanks.		
GEOGKEY	Number	TGN geographic place ID ("subject_id).		
		Created by MAP for MAPPLACES and		
		MAPSITES on new records		
PLACENAME	Text	Name of Place		
PLACENAMEFULL	Text	Concatenated String of place hierarchy, created		
		programmatically		
PLNAMEFULL_PLTYPE	Text	Concatenated String of place hierarchy		
		including place types, created programmatically		
PLTYPE	Text	Place Type (inhabited place, province, etc)		
PREFFLAG	Text	Preferred Flag - indicates "P"referred or		
		"V"ariant place name		
PLSOURCE	Text	Source for this record: TGN import,		
		MAPPLACE or MAPSITE.		
PLPARENT	Text	Place's Parent Place, next up on the hierarchy		
PARENTTYPE	Text	Parent Place Type		
PLPARENT_TERM_ID	Number	Parent's Place Term_ID (=PlaceANameID in		
		Doc Sources)		
PLPARENT_SUBJECT_ID	Number	Parent's Subject_ID (=GEOGKEY in Doc		
		Sources)		
PLPARENT_PLACEALLID	Number	Parent's Unique MAP ID		
GPARENT	Text	Place Grandparent Name		

GPTYPE	Text	Place Grandparent Place Type
GGP	Text	Place Great-Grandparent Name
GGPTYPE	Text	Place Great-Grandparent Place Type
GP2	Text	Place Great-Grandparent Name
GP2TYPE	Text	Place Great-Grandparent Type
RESID	Text	Researcher's ID Number
DATEENTERED	Date-time	Date Stamp for entries
PLACESMEMO	Memo	MAP Researcher Notes; not for public
ADDLRES	Yes/No	Additional Research Required
Term_Accent	Text	Term with accented characters
Language	Number	Not implemented.
OTHER_FLAGS	Text	TGN Field: indicates records displaying ISO
		and other international codes as placename
GEOGKEY_CHILDREN	Memo	Comma-separated list of all of the subregions
		of the current row; populated via UTL
		migration script

4.3 New database tables

The development team is going to create new database tables to accomplish the following functions: user action tracking [see 2.2.1]; revision of the database items [see 2.2.2]; community comments [see 2.5.1]; personal messages [see 2.5.3]; digitized manuscripts [see 2.7].

4.3.1 User action tracking

tblActions				
Field	Type	Description		
Timestamp	Date/time (automatic)	Date and time of creation		
Ip	Varchar	Ip address of the user		
UserID	Numeric	Extern Key		
Type	Text	Type of information (Action)		
Descr	Text	Description of information depending on the user action.		

The Chronology dropdown menu [see 3.2.1.5] and the Shortcut Menu [see 3.2.1.9] are created getting the data from this table.

4.3.2 Database Items revisions

Four new tables are created *tblRdocuments, tblRvolumes, tblRpeople, tblRplaces,* with the same structure of the *tblDocuments, tlbVolumes, tblPeople and tblPlaces* [see 4.2], but with three additional fields:

- realID: that is the current record
- timestamp: the timestamp of the item revision
- ID-user: a key or a username that identifies the person that modified that record

For example, a new record in the *tblRvolumes* with *realID* set as 1171a, *timestamp* set as 11122012, and *ID-user* set as lallori, is an old revision of the current 1171a, which has been modified by the user lallori on October 12 2012.

4.3.3 Community Comments

TblComments			
Field	Type	Description	
ID	Numeric (automatic)	UniqueID	
Timestamp	Date-time (automatic)	Date and time of creation	
UserID	Numeric	External Key (USERS)	
Туре	Numeric	0=Document, 1= Person, 2=Volume, 3=Place	
EntryID	Numeric	External Key (tblDocuments)	
Status	Numeric	0=TOBE_EVALUATED, 1=APPROVED,	
		2=NOT_APPROVED	
Text	Text	Comment text	

4.3.4 Personal Messages

TblMessages				
Field	Type	Description		
ID	Numeric (automatic)	UniqueID		
Timestamp	Date-time (automatic)	Date and time of creation		
UserID	Numeric	External Key (USERS)		
Туре	Numeric	0=Personal Message, 1= System Message		
Status	Numeric	0=READ, 1=NOT_READ		
Text	Text	Message text		

4.3.5 Showing digitized manuscripts

The digitized manuscripts reside on the servers filesystem as files. In the database tables a field is reserved to the images path.

TblDigitalDocs				
Id	Unique ID			
Vol_id	Text	Identificativo del volume		
Progr	Numeric	Numero progressivo del documento		
		digitalizzato		
Туре	Text	COPERTA=P, Carta=C, Guardia=G,		
		Allegato=A		
Folio	Numeric	Numero del foglio		

4.3.6 Google Maps integration

In the table *tblPlaces* a new field, *geocoord*, will be added. It will contain the geographic coordinates that will be used by the Google Maps APIs [see 6.7].

4.3.7 Catalog descriptions

A new table for catalog descriptions will be created, *tblCatalog*, using the fields listed in *Minutes of Meeting 3 – Catalog Description*.

5. Help System Design

5.1 The User Manual

In each user's profile page [see Image in section 2.5.2], the scholar is provided with the user manual. It will be available both in html and pdf versions. From this page, the user can also access the help videos section. The help videos are a step-by-step guide to perform some of the main activities a user performs during the day-to-day work:

- Data Entry System
 - Create a document file
 - Create a people file
 - Create a volume file
 - Create a location file
- How to perform an Advanced Search
 - On documents
 - On people
 - On volumes
 - On locations
 - On community comments
- How to deal with the community comments
- How to use the personal messaging tool
- How to deal with your user profile

During the Community Testing phase [see Appendix C – Preliminary Gantt Chart], the community testers will be able to ask the MAP team to create additional videos to guide them through tasks, which need further explanation, and are not mentioned above.

5.2 The embedded Help System

Aside from the User Manual, the Help System will be embedded inside the system, and will be accessed through the "?" icons.

Each "?" icon will open a help pop-up explaining that particular field or function.

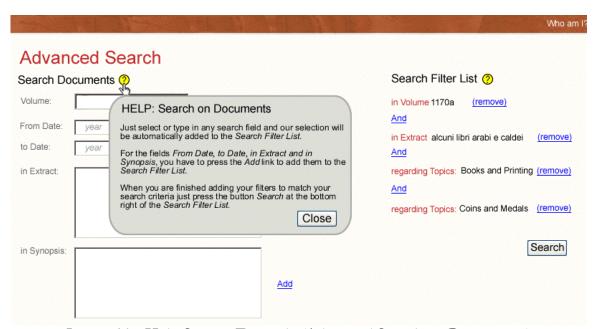


Image 46 – Help System Example (Advanced Search on Documents)

5.3 The Developer Manual

MAP is going to release the code produced as open source, which, therefore, will be used and adapted for other purposes by other institutions and/or individuals. For this reason, MAP must provide a detailed technical documentation.

The developer manual produced will consists in the Software's UML schemas (Class Diagram, Composite Structure Diagram, Sequence Diagram, and Activity Diagram) as well as the APIs documentation.

MAP Development team, while developing software code, has been documenting each java class produced. By using the Javadoc tool, which that is provided by the JDK, we will able to easily generate the APIs documentation in HTML format.

6. Architectural Strategies

Various web-oriented programming languages are available nowadays to create successful web interactive platforms. Excluding the Microsoft world, the software architects' attention focused on both PHP and JAVA. Comparing advantages and disadvantages of these two technologies, which are both very powerful and mature, MAP chose JAVA.

Therefore, the code will be written following the JEE specifications.

6.1 Frameworks

6.1.1 Spring MVC

The development team will use the *Spring MVC*, which is a Java framework to build web applications on top of the Java Enterprise Platform. Spring MVC architecture meets MAP's three-layer design methodology.

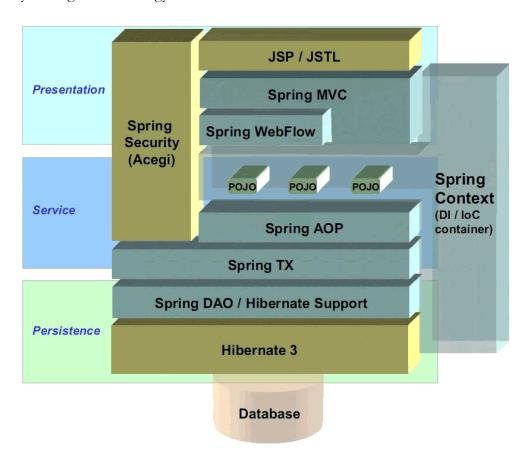


Image 47 – Spring Architecture

The Developers will also use the *DWR* library (*Direct Web Remoting*), also called *Easy Ajax for Java*, which provides an easy and fast interaction between Java Server Pages and JavaScript client language.

6.1.2 Security and Authentication

To develop the authentication and authorization mechanisms the development team will use the *Spring Security* framework. This provides a set of modules to connect to the most important authorization and authentication mechanisms, such as *OpenLDAP*, which will be used to store username and passwords as well as the profile fields data.

This security framework is an ad-hoc module completely separate from the application, which manages the users' authentication and the authorization.

Thus, it is not necessary to store the username and password in the database tables and most importantly, during the development of the software, the developers do not have to worry about introducing control methods on the users' authentication: this process is entirely delegated to the *Spring Security* framework.

6.1.3 Data Model

An object-relational mapping type (ORM) approach to the data model will be used.

To accomplish these tasks, the development team will use the Java Persistence API (JPA) linked to the Hibernate implementation. Hibernate will be configured for Mysql. In addition, the Java Persistence Query Language with its sql-like syntax will help the developers to deal with the database queries more easily and to cut down the time of developing the Advanced Search sub-module.

This model does not bind us to this database, making any migration to another rdbms fast and easy.

6.1.4 System Logs

For the system logs the framework Apache Log4J will be used. It will be configured to append the data to Mysql as well as to text log files.

6.1.5 Graphics

The *Spring MVC* framework, allows us also to link the User Interface Layer to the *Apache Tiles 2* API. The Web Graphic Designer will use this 2 API, to produce highly reliable, usable and portable W3C-compliant HTML templates.

JQuery will be used along with the DWR to create Javascript widgets for the auto complete text fields and the auto complete dropdown suggestion boxes.

6.1.6 Google Maps integration

To integrate the Google Maps in the GeoBase, the official Google Maps API, provided by Google Code, will be used. These APIs are released under the Google Maps/Google Earth APIs terms of service.

6.2 Database

MAP's current online read-only application (http://documents.medici.org) is using as data source the *Mysql 5.x*. This database is already up and running since April 2006 and did not have any kind of failure and/or performance problems. For these reasons MAP chose to continue using *Mysql 5.x*; this will provide 90% of the database structure ready to be used for the new system.

To migrate the data from Microsoft Access to Mysql, MAP now uses some *Perl-based* scripts, which will need some minor changes to be used with the new system. These will be used to populate the database with data for the testing phase, and also for the final transition from the old system to the new one.

6.3 Software Project Management and development tools

The open source software project management tool *Apache Maven* will be used to manage the development of the software. It allows developers in a team-working environment to deal easily with code compiling and distribution.

The *Eclipse IDE* will be used, in conjunction with the above-mentioned *Apache Maven*, for building and testing the web-application.

6.4 Application Server

Even if it is MAP's intention to use *Tomcat* (which is a servlet container), if the development team starts from the beginning to follow the JEE specifications, this will permit MAP to easily migrate to the use of *JBoss* in the future, if needed.

6.5 Bug-tracking

For the testing phase, (and also when in production) MAP and the test users will use the *Bugzilla* bug-tracking system. This system will help to keep track of new bugs in the development; this will be used also to avoid using emails to receive bugs notifications.

6.6 Server-side scripting

Bash scripting will be used for both backup purposes and to prepare the digitized images for the use with IIPImage (Manuscript Viewer).

7. Application Environments

7.1 The Development Environment

The Development Environment consists in the platform where the Java developers and the Web Graphic Designer work.

The two Java developers have a Microsoft Windows or Linux-based laptop (depending on their preference) with these services installed:

- Apache Tomcat 6.0.x
- Mysql Client
- Jsdk 1.6
- Eclipse IDE for Java EE Developers

The Web Graphic Designer has a Microsoft Windows-based laptop with these services installed:

- Apache Tomcat 6.0.x
- Mysql Client
- Jsdk 1.6
- Eclipse IDE for Java EE Developers
- Photoshop or Macromedia Studio (depending on the Web Graphic Designer's preference)

7.2 The Testing Environment

The Testing Environment resides on one server in MAP's room in the State Archive in Florence.

It runs Linux Slackware and these services will be running:

- Apache Tomcat 6.0.x
- Mysqld 5.x
- Mysql Client

7.3 The Production Environment

The production environment resides on MAP's outsourced servers [see 8.1].

It runs Linux CentOS and these services will be running:

- Apache Tomcat 6.0.x
- Mysqld 5.x
- Mysql Client
- Subversion or CVS
- OpenSSH
- Bugzilla

8. Server System

8.1 Servers deployment

One of the main bottlenecks of a web-based system that has to serve multiple users simultaneously is the Internet bandwidth. For this reason MAP chose an outsourced solution instead of an in-house one.

MAP is going to rent two dedicated private servers with the following specifications:

- Intel Xeon (4x core, 1,86Ghz)
- 2Gb Ram
- 2 x 300Gb SAS Hard Drives (RAID 1 configuration)
- 10Mbps assured internet bandwidth
- Public IP

The two hard drives on each server create a RAID 1 array (mirroring) and are managed by a hardware raid controller.

MAP's system requires continuous availability and a high degree of reliability. For this reason, one server will act as a master server, the other one as a slave; all the data will be synchronized. If a failure occurs, the slave server takes the lead, providing a *failure* capability [see 8.1.1 and 8.1.2].

8.1.1 Servlet container: Tomcat configuration

Each server will run an instance of Tomcat. The two Tomcats will be configured for a high availability using Tomcat's 6 cluster capabilities. A *watchdog-like* component is included in software architecture; this will act as a real time monitor for the software. In case of system malfunction the request will be forwarded to the other Tomcat instance, so that the user will not lose requests and/or sessions. Administrators will be notified of any system malfunction by the watchdog.

8.1.2 Database: Mysql configuration

As regards the database, when one Mysql instance fails, the service must be provided through another route. Therefore we will need process redundancy. To gain this we will install and configure Mysql on both servers building an Active/Standby Mysql configuration.

Only one Mysql instance will be active, and when a failure is detected the standby instance will be made active.

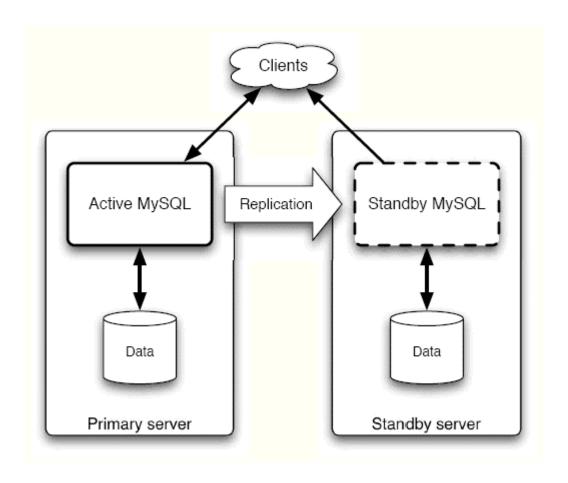


Image 48 – Active/Standby Mysql configuration

The two instances will be synchronized using scheduled time. These times will be scheduled during the part of the day in which there is less workload: this will be determined and adjusted using the system statistics capabilities. When the standby instance is activated, an additional synchronization process is started.

8.1.3 Backup and Restore Procedures

Backups are managed through the system Administrative Interface [see 2.6.3]. Scheduled backups automatically back up the system, but manual backups are also available.

The restore procedure is accessible from the web graphic interface, but also an automated one from the operating system shell will be implemented in case of critical failure.

8.1.4 Hardware and bandwidth upgrading

Even though MAP considers this hardware configuration and the dedicated Internet bandwidth proportioned to the expected usage (or workload) of the web-interface, the application software and the operating system software that compose the new MAP system are structured to permit easy migration to a hardware configuration, which provides a higher performance such as cluster or cloud solutions. Moreover, the web hosting company we are going to rely on granted us prompt access to a wider Internet bandwidth, upon request.